DEUTSCHE
BUNDESBANK

PPP: a Disaggregated View

Christoph Fischer

Discussion Paper
Series 1: Studies of the Economic Research Centre
No 07/2004

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-1
Telex within Germany 41227, telex from abroad 414431, fax +49 69 5601071

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax No +49 69 9566-3077

ISBN 3–935821–85–9

Abstract

By disaggregating price indices, it becomes apparent that the real exchange rate consists of the real exchange rate for a single good and a weighted sum of relative prices between goods. When applying a battery of panel unit root tests to this sum and its components, it is found that both the sum and the relative prices are non-stationary. This implies that PPP is invalid even if the LOP holds for all goods. The findings contrast with the result from panel unit root tests that real exchange rates as a whole are stationary. Several suggestions for solving the conflict are discussed.

JEL classification: F31, C33

Keywords: purchasing power parity, real exchange rate, panel unit root tests

Non-technical summary

In the economic policy debate on the appropriateness of exchange rates, notions of the respective equilibrium value often serve as a reference variable. Purchasing power parity (PPP) is the simplest and most popular concept for determining an equilibrium exchange rate. In a plethora of published papers, however, there is still a debate on whether or not the PPP hypothesis is supported by the data. From a theoretical point of view, there are at least two reasons why, even in the long run, PPP may not hold. The first of these is exploited in the hypothesis of Balassa (1964) and Samuelson (1964): If some goods are internationally non-tradable, PPP does not necessarily hold because there is, by definition, no arbitrage between non-tradable goods.

This paper examines a second, old but rarely mentioned precondition for PPP to hold. As Devereux (1997) puts it, if "... the composition of price indices differs across countries, ... trend movements in relative goods prices will lead to persistent deviations from PPP". Until now, there does not appear to have been a systematic empirical examination of whether or not this is the case. In order to formalize the argument, economy-wide price indices have been disaggregated into price indices for single goods. It becomes apparent that the real exchange rate consists of two components; first, the real exchange rate for a single good and, second, a weighted sum of relative prices between *different* goods. PPP requires that the real exchange rate as a whole is stationary. Since stationarity of the first component, the real exchange rate of a single good, implies that the law of one price holds, the second component should be stationary as well in order to maintain the validity of PPP.

Using a battery of panel unit root tests, it is first shown that relative prices between different goods are predominantly non-stationary. Devereux's (1997) trend movements in relative goods prices do exist. In a second step, it is shown that this does not change if weighted relative prices are related to their foreign counterparts. This could be a consequence of differences in the composition of price indices across countries. In a third step, it is found that the non-stationarity does not aggregate out. The second component as a whole is non-stationary.

The results imply that the investigated condition for PPP to hold is not fulfilled. While the results are consistent among themselves, they are at odds with the finding of many papers that panel unit root tests suggest the stationarity of real exchange rate as a whole – a result that has been confirmed here. Two ways of solving the conundrum are proposed. Since it is found that the (variance in the difference of) the non-stationary second component is small compared with the possibly stationary first component, the latter could simply mask the former. This has been suggested by Bayoumi/MacDonald (1999) in a related context. Alternatively, the fact that panel unit root tests find that real exchange rates are mean-reverting could be due to a bias. This bias arises if variables are tested that consist of one stationary and one non-stationary component, which is exactly what this paper suggests if the law of one price holds. In this respect, the paper supports Engel's (2000) suggestion that unit root tests are subject to such a size bias when they are applied to real exchange rates. In contrast to Engel (2000), who derives this result in a Balassa-Samuelson framework, however, it is found here that this follows even if the law of one price holds for every single good.

Nichttechnische Zusammenfassung

In der wirtschaftspolitischen Diskussion über die Angemessenheit von Wechselkursen dienen Vorstellungen über den jeweiligen Gleichgewichtswert häufig als Referenzgröße. Die Kaufkraftparitätentheorie ist das einfachste und populärste Konzept zur Bestimmung eines gleichgewichtigen Wechselkurses. Es wird allerdings in einer Vielzahl von Publikationen immer noch darüber diskutiert, ob die Kaufkraftparitätenhypothese empirisch unterstützt wird oder nicht. Aus theoretischer Sicht gibt es mindestens zwei Gründe, warum die Kaufkraftparitätentheorie auch in der langen Frist nicht zu gelten braucht. Der erste von ihnen wird in der Hypothese von Balassa (1964) und Samuelson (1964) verwendet: Wenn einige Güter international nicht-handelbar sind, braucht die Kaufkraftparitätentheorie nicht zu gelten, weil es definitionsgemäß keine Arbitrage zwischen nicht-handelbaren Gütern gibt.

Im vorliegenden Diskussionspapier wird eine zweite Bedingung für die Geltung der Kaufkraftparitätentheorie untersucht. Diese Bedingung ist zwar schon lange bekannt, wird aber selten erwähnt. Ein Beispiel ist Devereux (1997): Wenn „ ... sich die Zusammensetzung des Preisindex in einem Land von dem eines anderen Landes unterscheidet, werden Trends bei relativen Güterpreisen zu dauerhaften Abweichungen von der Kaufkraftparitätentheorie führen.“ Bis heute scheint eine systematische empirische Untersuchung darüber zu fehlen, ob dies der Fall ist oder nicht. Um das Argument zu formalisieren, wurden gesamtwirtschaftliche Preisindizes in Preisindizes für einzelne Güter disaggregiert. Auf diese Weise stellt sich heraus, dass der reale Wechselkurs aus zwei Komponenten besteht: erstens aus dem realen Wechselkurs für ein einzelnes Gut und zweitens aus einer Summe gewichteter Relativpreise zwischen *verschiedenen* Gütern. Die Kaufkraftparitätentheorie erfordert, dass der reale Wechselkurs als Ganzes stationär ist. Weil die Stationarität der ersten Komponente, des realen Wechselkurses eines einzelnen Gutes, impliziert, dass das Gesetz der Unterschiedslosigkeit der Preise gilt, sollte auch die zweite erwähnte Komponente stationär sein, damit die Kaufkraftparitätentheorie erfüllt ist.

Mit Hilfe mehrerer Panel-Einheitswurzeltests wird zunächst gezeigt, dass Relativpreise zwischen verschiedenen Gütern weitgehend nicht-stationär sind. Relative Güterpreise weisen in der Tat die von Devereux (1997) beschriebenen Trends auf. In einem zweiten Schritt wird ermittelt, dass sich dies auch nicht ändert, wenn gewichtete Relativpreise im Inland zu entsprechenden ausländischen Zeitreihen in Beziehung gesetzt werden. Dies könnte auf Unterschiede in der Zusammensetzung der Preisindizes verschiedener Länder zurückzuführen sein. In einem dritten Schritt wird festgestellt, dass auch eine Aggregation über alle Güter hinweg die Nicht-Stationarität nicht beseitigt. Die zweite Komponente als ganze ist also nicht-stationär.

Die Ergebnisse implizieren, dass die oben genannte Bedingung für die Gültigkeit der Kaufkraftparitätentheorie nicht erfüllt ist. Diese Ergebnisse sind zwar untereinander konsistent, widersprechen aber den Resultaten zahlreicher Studien, denenzufolge Panel-Einheitswurzeltests Stationarität des realen Wechselkurses als ganzem nahe legen. Letzteres wird auch im vorliegenden Diskussionspapier bestätigt. Es werden zwei mögliche Lösungen für den scheinbaren Widerspruch vorgeschlagen: Da es sich herausstellt, dass die nicht-stationäre zweite Komponente (gemessen an der Varianz ihrer Differenz) verglichen mit der möglicherweise stationären ersten Komponente klein ist, könnte diese zweite Komponente leicht von der ersten überlagert worden sein. Diese Auffassung vertreten Bayoumi/MacDonald (1999) in einem ähnlichen Zusammenhang.

Alternativ könnte die Tatsache, dass sich der reale Wechselkurs bei Verwendung von Panel-Einheitswurzeltests als stationär herausstellt, auf eine Verzerrung zurückgeführt werden. Eine solche Verzerrung tritt auf, wenn Variablen getestet werden, die aus einer stationären sowie aus einer nicht-stationären Komponente bestehen. In der Tat ergibt sich den Ergebnissen des vorliegenden Diskussionspapiers zufolge eine solche Konstellation unter der Voraussetzung, dass das Gesetz der Unterschiedslosigkeit der Preise gilt. Auf diese Weise bekräftigt das vorliegende Papier Engels (2000) Ansicht, nach der Einheitswurzeltests einer solchen Verzerrung unterliegen, wenn mit ihnen reale Wechselkurse getestet werden. Im Unterschied zu Engel (2000), der dieses Ergebnis im Rahmen eines Balassa-Samuelson-Ansatzes ableitet, ergibt es sich hier selbst dann, wenn das Gesetz der Unterschiedslosigkeit der Preise für jedes einzelne Gut gilt.

Table of Contents

List of Tables and Figures

Tables

Figures

PPP: a Disaggregated View*

1. Introduction

Since the late 1980s, there has been a plethora of published papers in which unit root or cointegration tests are used to check the validity of purchasing power parity (PPP). This literature is surveyed in the review papers of Breuer (1994), Froot/Rogoff (1995), Rogoff (1996), Lan (2002), and Sarno/Taylor (2002), ch. 3. In the first few years, it was mainly *univariate* unit root or cointegration tests that were applied. What has emerged, however, is that univariate unit root tests have relatively low power when applied – as has commonly been the case – to observation periods spanning, for instance, the post-Bretton Woods era. Therefore, *panel* unit root tests have regularly been used instead in more recent papers which investigate PPP during the post-Bretton Woods era. The gain in information through pooling of data, it is argued, compensates for the problem of low power. Some recent panel unit root studies of PPP are Taylor/Sarno (1998), Higgins/Zakrajšek (1999), Coakley/Fuertes (2000), Fleissig/Strauss (2000), Choi (2001), Kuo/Mikkola (2001), Papell/Theodorides (2001), Parsley/Popper (2001), Wu/Wu (2001), and Ho (2002), to name just a few. In contrast to most univariate unit root tests of PPP, the majority of panel unit root tests find evidence in favour of PPP, one notable exception being O'Connell (1998).

From a theoretical point of view, an empirical result supportive of PPP is not a matter of course. PPP (in its relative version) claims that there is a long-run tendency for the change in the nominal exchange rate of two countries to equal the difference between the inflation rates of their CPI, PPI or WPI baskets of goods. This claim is based on the law of one price (LOP), according to which the price of a single good in one country should tend to be the same as the price of the same good in another country if prices are expressed in a common currency. Arbitrage will prevent any substantial deviations. Aggregating over all goods in the economy should result in PPP (see Dornbusch, 1987, Obstfeld/Rogoff, 1996, p 202, or Sarno/Taylor, 2003, p 52). Apart from arbitrage and the LOP, there is, of course, a second mechanism which could maintain PPP: PPP will also hold if all the disturbances satisfy the conditions of the homogeneity postulate of monetary theory in the sense that they leave

* I would like to express my gratitude to Janice Boucher Breuer for providing me with her RATS program for simulating SURADF critical values. I would like to thank Jörg Breitung, Ulrich Grosch, Axel Jochem, Joachim Keller, Daniel Porath, Karin Radeck and seminar participants at the Bundesbank for their valuable suggestions and comments. All remaining errors are my own.

unchanged all equilibrium relative prices, and thus lead only to an equiproportionate change in money and all prices, including the price of foreign exchange (Dornbusch, 1987).

There are, however, at least two reasons why, even in the long run, PPP may not hold. The first of these is exploited in the hypothesis of Balassa (1964) and Samuelson (1964): If some of the goods in the respective baskets are internationally non-tradable, PPP does not necessarily hold because there is, by definition, no arbitrage between non-tradable goods. This means that, for these goods, there is no reason for the LOP to be fulfilled. DeLoach (2001), Canzoneri et al (1999), and De Gregorio et al (1994) are among the papers to find some evidence of Balassa-Samuelson effects in OECD countries. In a particularly interesting paper, Engel (2000) demonstrates how such results can be reconciled with evidence in favour of PPP. He argues that the Balassa-Samuelson modelling framework implies that the real exchange rate consists of a stationary and a non-stationary component. He shows that this two-component character severely biases unit root tests in favour of rejecting non-stationarity of the real exchange rate, and concludes that results supportive of PPP may be due merely to a misspecification of the data-generating process of real exchange rates.

Apart from the Balassa-Samuelson argument, there is a second, less well-known reason why PPP may not hold. If production and consumption patterns differ across countries, a given single good receives a weight in the CPI, PPI or WPI basket of country *i* which is different from the weight of the same good in the basket of country *j*. Thus, any change in the relative world market price between two different single goods affects the price indices of the two countries differently, even if the LOP holds continuously for every single good. An oil price shock, to take the most simple example, has a greater effect on the CPI of a country like Canada, which, proportionally, consumes a greater amount of oil, than it does on the CPI of most other countries. Apart from oil price shocks, every technological or demand shock that does not affect all goods equally causes relative prices between different goods to change. If the composition of price indices differs across countries, an aggregation of the LOP will, in the presence of such shocks, no longer result in PPP. Nor would the other mechanism for maintaining PPP work, obviously, because the change in relative prices would violate the homogeneity postulate of monetary theory.

This second argument against PPP is only occasionally mentioned in the literature. Devereux (1997), p 777, for instance, writes “But the composition of price indices differs across countries, so that trend movements in relative goods prices will lead to persistent deviations from PPP.” Sarno/Taylor (2002), pp 53-54, note also that differences in weights across countries may lead to deviations from PPP, if “... price impulses impinge heterogeneously across the various goods and services in an economy ...”. An earlier example of this literature is Hsieh (1982). Bayoumi/MacDonald (1999) find that, while international real exchange rates are stationary, real exchange rates across US regions are non-stationary, and explain this result in terms of heterogeneously acting price impulses which may be detectable at a regional level but may be swamped by homogeneous shocks at a national level.

Despite its acknowledged potentially detrimental effect on PPP, however, the issue of persistent changes in relative prices in combination with differing weights has, to my knowledge, not yet been investigated empirically in a systematic way, possibly owing to a lack of adequate data. The present paper fills this gap by exploiting a new OECD database. By applying different panel unit root tests, it investigates for a panel of OECD countries whether Devereux’s trend movements in relative goods prices can be detected at a sectoral level, whether they are still detectable when they are weighted and related across countries, whether they aggregate out over the entire economy, and what the repercussions on (tests of) PPP are.

To this end, in section 2 the real exchange rate will be broken down into its components that are mainly weighted relative goods prices, and it will become apparent that three alternative conditions on various aggregation levels of these components are sufficient for PPP to hold. After presenting the data in section 3, several panel unit root tests on the components of real exchange rates in different aggregations will be performed in section 4. This amounts to an examination of the derived conditions for the validity of PPP. It proves to be the case that, for all countries, there is strong evidence of most relative prices being non-stationary, that there is no indication of weights forming a cointegrating vector, and that this non-stationarity apparently does not aggregate out; ie none of the conditions is fulfilled.

Section 5 provides some thoughts on how these results can be reconciled with the evidence supporting the stationarity of real exchange rates as a whole. It becomes evident, in particular, that the stochastic trend in relative goods prices imparts exactly the two-component structure to the real exchange rate that Engel (2000) derived in a Balassa-Samuelson framework; in other words, even if there are no Balassa-Samuelson effects because the LOP holds continuously for every single good, the real exchange rate consists of one stationary and one non-stationary component. This result supports Engel’s (2000) conjecture that unit root tests of real exchange rates are biased in favour of rejecting non-stationarity. A second explanation for the conflicting results is offered: Since the variance in the difference of the non-stationary component is relatively small compared with the possibly stationary one, the former might be obscured by the latter. Thus, the results also support the suggestions made by Bayoumi/MacDonald (1999).

2. Conditions for the validity of PPP

The (absolute) PPP hypothesis may be expressed as

$$(1) \qquad S_{it} = \frac{P_{it}}{P_{jt}},$$

where S_{it} is the price of the numéraire country j's currency expressed in units of country i's currency at time t, P_{it} (P_{jt}) is the price of a basket of goods consumed or produced in country i (j). Denoting the log of a variable with a small letter, the log of the bilateral real exchange rate between country i and the numéraire country is defined as

$$(2) \qquad r_{it} \equiv s_{it} - p_{it} + p_{jt}.$$

In the typical empirical application, some national price indices, for instance CPIs, are used to calculate r_{it}, which is tested for stationarity afterwards. Stationary real exchange rates may be interpreted as evidence in favour of PPP.[1]

In order to obtain components of the real exchange rate, the price indices must be disaggregated. If they are calculated as geometric indices, each p_{it} is the weighted sum of the (log of the) prices of all the individual goods $k = 0, \ldots , m$ in country i at time t, p_{ikt},

$$(3) \qquad p_{it} \equiv \sum_{k=0}^{m} \alpha_{ik} p_{ikt}$$

$$= p_{i0t} + \sum_{k=1}^{m} \alpha_{ik} \left(p_{ikt} - p_{i0t} \right),$$

where α_{ik} is the weight of good k in country i, $\sum_{k=0}^{m} \alpha_{ik} = 1$ for all i, and an arbitrary $k = 0$ is the numéraire good. Inserting (3) into (2), and using for $k = 0$

$$(4) \qquad r_{ikt} \equiv s_{it} - p_{ikt} + p_{jkt},$$

[1] Actually, (1) additionally requires $r_{it} = 0$ in the long run, which is the condition that delimits the absolute version of PPP from relative PPP. However, only relative PPP is usually examined because it is generally maintained that tariffs and transaction costs result in $r_{it} \neq 0$. Furthermore, if prices are measured by indices, r_{it} may be different from zero.

which, analogously to (2), defines the relative price of the individual good k *across* countries expressed in a common currency as r_{ikt}, yields

$$(5) \qquad r_{it} = r_{i0t} + \sum_{k=1}^{m} \left[\alpha_{jk} \left(p_{jkt} - p_{j0t} \right) - \alpha_{ik} \left(p_{ikt} - p_{i0t} \right) \right].$$

According to (5), the real exchange rate consists of two components; first, the real exchange rate for a single good $k = 0$ and, second, a weighted sum of relative prices between *different* goods. The time series properties of the terms on the right-hand side of (5) should determine whether or not tests of PPP are rejected. Stationarity of the first term, r_{i0t}, implies that the LOP holds for the numéraire good. For the time being, let us consider a best-case scenario for PPP by assuming stationarity of r_{i0t}. The issue of the validity of the LOP has already been analyzed in previous studies and will be taken up again in section 5.

Equation (5) reveals, however, that the stationarity assumption on r_{i0t}, is not sufficient for PPP to hold. Each of the other terms on the right-hand side of (5) consists of the relative price between a good k and the numéraire good *within* one country multiplied by the respective weight of good k in the price index of this country. With regard to these components of the real exchange rate, one of the following conditions has to be fulfilled for PPP to hold (always assuming that the LOP holds for $k = 0$) since, otherwise, the real exchange rate r_{it} would not be stationary.

Condition 1: If *all* the relative prices between every good k and the numéraire good within both country i and the numéraire country are stationary, PPP will hold. A violation of this condition would imply that Devereux's (1997) trend movements in relative goods prices are present. This could occur as a consequence of significant technological or demand shocks that do not affect all goods equally.

Condition 2: Given that condition 1 is not fulfilled, PPP holds nevertheless if *all* the m differences between such a relative price in country i multiplied by its respective weight in country i's price index, on the one hand, and the same variable in the numéraire country j, on the other (the expressions in square brackets in (5)) are stationary. This would occur if, for instance, owing to similar technology and demand patterns or just because of open trade, the relative prices behave similarly in both countries ($(p_{jkt} - p_{j0t}) \approx (p_{ikt} - p_{i0t})$) and, at the same time, the weights of the individual goods do not differ much across countries ($\alpha_{jk} \approx \alpha_{ik}$).

Condition 3: Given, finally, that neither condition 1 nor condition 2 is fulfilled, PPP still holds, if the non-stationarity disappears when aggregating across goods, that is if the individual non-stationary components cointegrate for some reason so that the second-term sum in (5) becomes stationary.

How would a violation of these conditions impinge on the mechanisms that are supposed to maintain PPP? By definition, if condition 1 is rejected, the mechanism which relies on the sole existence of disturbances that leave relative prices unchanged does not work. While an examination of conditions 1-3 does not deal explicitly with the other mechanism that is assumed to maintain PPP – arbitrage and the LOP – this mechanism cannot re-establish PPP if the stationarity of the second-term sum (condition 3) is rejected. Even if the LOP holds for every single good and r_{ikt} is therefore stationary for every k, this only implies that the first term on the right-hand side of (5) is stationary for an arbitrary numéraire good k. For the real exchange rate to be stationary and thus for PPP to hold, it is additionally required that the second-term sum in (5) is stationary as well.

In order to provide a deeper understanding of PPP, the three conditions will be systematically investigated in the following sections by applying several panel unit root tests.

3. The data

The database that will be used below is the new "Structural Analysis Database" (STAN) provided by the OECD. This database not only includes but also extends the OECD's former "International Sectoral Database" (ISDB), which has been used inter alia by Wei/Parsley (1995), Engel (1999) and Sarno et al (2004). STAN, in principle, comprises data on all the sectors (goods and services) of all OECD countries. The sectoral division is based on the International Standard Industrial Classification (ISIC) Rev. 3. STAN does not, of course, provide data at an individual goods level. Obviously, it would have been convenient to work with price data that do not apply to baskets of goods. However, that is no reason to refrain from performing the analysis at a sectoral level,[2] especially if it can be shown that, even at this level, there are severe problems in meeting the criteria for PPP to hold. Furthermore, more disaggregated data were simply unavailable for a panel of countries.

At the chosen aggregation level, the total economy comprises 18 sectors. They are listed in the appendix. STAN provides annual data from 1977 to 1999 for *all* the 18 sectors of 11

---

[2] As, for example, the paper by Sarno et al (2004) shows, it is not uncommon in econometric applications to replace goods by sectors.

OECD countries, ie Austria, Belgium, Canada, Denmark, Finland, Germany, Italy, Japan, (South) Korea, the United Kingdom, and the USA. Disaggregating the sectors further or extending the observation period or the number of countries would have involved missing values in at least some of the panels. Note, however, that we need data on all sectors of the economy for an examination of condition 3. The annual frequency is relatively low and, therefore, only a limited number of observations is available. As is well known, however, a higher frequency leads to no more than a small increase in the power of unit root tests (see, for example, Campbell/Perron, 1991). Low frequency will therefore not be a serious problem, especially as data from a panel of time series will be combined throughout.

Like the earlier ISDB, STAN provides time series for value added at current prices and a quantity index for value added at constant prices. As in Wei/Parsley (1995) and Sarno et al (2004), value added at current prices has been divided by value added at constant prices in order to obtain (producer) price indices for each of the 18 sectors in each of the 11 countries. The price index series have been rebased so as to set the value of 1995 to 100. Weights have been computed by dividing value added at current prices in sector *k* by value added at current prices in the total economy which is also provided by STAN. Note that these weights are not necessarily constant over time.

As is common practice in the literature, the USA and, alternatively, Germany have been taken as the numéraire country. In order to obtain relative prices, price indices of each sector *k* have been divided by the price index of a numéraire *sector*, and logs have been taken. By analogy with taking the biggest economy, the USA, as numéraire country, the private sector which had, on average, the largest weight in the economy (ie the "Finance, insurance, real estate and business services" sector) was chosen as the numéraire sector. Note that the choice of the numéraire sector is independent of the discussion about tradability because it is not the Balassa-Samuelson argument which is considered. Trends in domestic relative prices as well as differences in production patterns across countries may occur for both tradables and non-tradables sectors. As a robustness check, however, a second set of relative prices has been constructed by using as numéraire sector the biggest sector that is usually classified as consisting of tradables: "Machinery and equipment". Results relating to this set of relative prices are presented in section 4.3.

In order to give an impression of the data used, Figure 1 shows the relative prices for all sectors in one of the (numéraire) countries, the USA. Figure 1 reveals that, while the relative price of some sectors such as "Food products, beverages and tobacco", "Pulp, paper, paper products, printing and publishing", "Transport equipment", or "Community social and personal services" has tended to remain constant in comparison with the numéraire sector, most relative prices exhibit a tendency to decline. This is preliminary evidence against condition 1, which requires that all relative prices should be stationary. Since it is not sufficient to find only one or some of the relative prices to be stationary, this negative preliminary result would have been found irrespective of the choice of numéraire sector.

For the computation of real exchange rates as defined in (2), nominal exchange rate data are taken from the IMF's International Financial Statistics, line ae, and the price indices for the total economy are again taken from the OECD's STAN.

The data thus allow the usual country panels to be formed in order to test for stationarity of the real exchange rate. These panels comprise ten real exchange rate series with 23 annual observations each for the period 1977-1999. In addition, however, panels of 17 domestic relative price series can be formed, one for each of the 11 countries. They will be used to check for condition 1. Similarly, using the calculated weights, panels of the weighted ratio of domestic and numéraire country relative prices with N = 17 can be constructed, which will allow an examination of condition 2. Since data on all sectors are available, these variables can be aggregated across sectors, yielding an expression for the second-term sum in (5). Since this can be done for each country, a panel of ten series of sums, one for each of the countries, will result. This panel can be used to check for condition 3. All of the panels that will be examined are balanced panels with T = 23.

4. Unit root tests of the real exchange rate and its components

In this section, panel unit root tests will be used first to replicate the results from the literature concerning stationarity of real exchange rates and then to test successively for the three conditions, each of which is sufficient for PPP to hold.

4.1 Methods

For the examination of each hypothesis, four unit root tests will be applied, the Levin and Lin test (referred to below as LL) developed by Levin et al (2002), the t-bar test (referred to below as IPS) proposed by Im et al (2003), Taylor/Sarno's (1998) MADF test, and the SURADF test of Breuer et al (2001, 2002). All these tests in some way consider the N-equation model

(6) $$\Delta y_{i,t} = \mu_i + \rho_i y_{i,t-1} + \delta_{i,1}\Delta y_{i,t-1} + \delta_{i,2}\Delta y_{i,t-2} + \ldots + \delta_{i,p_i}\Delta y_{i,t-p_i} + \varepsilon_{i,t}$$

where $y_{i,t}$ is the variable to be tested for stationarity, $i = 1, 2, \ldots, N$ denotes the country or sector, $t = 1, 2, \ldots, T$ is time, the error term $\varepsilon_{i,t}$ is white noise but correlations across individuals are allowed in the MADF and the SURADF tests, and p_i is the maximum lag in the equation for $y_{i,t}$. As is usual in testing for PPP, a linear deterministic time trend has not been included in (6) for any of the panels, first, because it is inconsistent with PPP (see, for example, Higgins/Zakrajšek, 1999, Kuo/Mikkola, 2001, Papell/Theodorides, 2001) and, second, because preliminary single-equation ADF tests provide no evidence of significant deterministic trend components according to the critical values of Dickey/Fuller (1981), p 1062.

The LL test and the IPS test have been chosen because they would appear to have become established as standard panel unit root tests and are often used as benchmark. Both of these tests ignore correlations of the residuals across individuals. In order to allow for at least a limited degree of dependence, time-specific intercepts have been included on the right-hand side of (6) in both tests by subtracting cross-sectional averages from all observations.[3] The lag length p_i has been determined for both tests by successive elimination of the greatest insignificant lag in a univariate ADF equation for each variable $y_{i,t}$, the maximum possible lag generously being six years. In most cases, however, one lag proved to be sufficient. Without going into the details of these rather standard tests, it is important to note that LL tests for the null hypothesis H_0: $\rho_i = 0$ for all i against the alternative H_A: $\rho_i = \rho < 0$ for all i. In most economic applications, for instance, when testing for PPP, the equality restriction under the alternative is not particularly plausible. Furthermore, the possibility that some of the variables are stationary and some are not is not dealt with. In these respects, the IPS test is preferable because it tests the null hypothesis H_0: $\rho_i = 0$ for all i against the alternative H_A: $\rho_i < 0$ for $i = 1, \ldots, N_1$ and $\rho_i = 0$ for $i = N_1+1, \ldots, N$.

Both the LL and the IPS tests have been criticized because they do not take correlations across individuals into account. Time effects are a weak substitute at best (see, for example, O'Connell, 1998, Strauss/Yigit, 2003). When testing for stationarity of real exchange rates, this cannot be ignored because cross-correlations will usually be high. Therefore, it is often suggested that explicit account be taken of the covariance matrix of the residuals in a SURE framework. Here, Breusch/Pagan's (1980) Lagrange multiplier tests on the diagonal form of this matrix have tentatively been performed. The diagonal form has indeed been rejected not only for the real exchange rate panels but for nearly all other panels as well. Cross-correlations should therefore not be ignored when testing for stationarity. As far as relative prices are concerned, this result may be due, for instance, to

---

[3] All the LL and IPS tests have also been performed without such time effects. The results of these tests (which are available on request) did not, however, differ much from the ones with time effects.

technological or oil price shocks that change domestic relative prices in more than one sector simultaneously.

In order to take account of such cross-correlations, the multivariate ADF test of Sarno/Taylor (1998) and Taylor/Sarno (1998) – MADF in short – has been used. It is based on a SUR estimation of the N-equation system (6) without any further restriction. A Wald test statistic is computed on the combined restrictions of the null hypothesis H_0: $\rho_i = 0$ for all i. The critical values are computed by simulation. It is important to note that the response surface estimation of the critical values which is suggested in Taylor/Sarno (1998) cannot be used for panels of the size examined here, which is either $N = 10$ and $T = 23$ or $N = 17$ and $T = 23$. In the former case, the simulated critical values are two to three times as large as the ones computed via response surface estimation, and in the latter case, they are even ten times as large.[4]

The null and alternative hypotheses of the IPS and the MADF tests are the same. Unfortunately, these hypotheses are not entirely adequate for testing conditions 1 and 2. The validity of these conditions depends on whether *all* the series of a panel are stationary. The null hypothesis of the IPS and MADF test is, however, rejected if there is at least one stationary series in the panel. Thus, rejections should occur even if only one of the series in the panel is stationary, although this is clearly not sufficient for fulfilling conditions 1 or 2. Sarno/Taylor's (1998) and Taylor/Sarno's (1998) JLR statistic is one of the panel unit root tests specifically designed to deal with these problems. Here, however, it cannot be applied because it implies estimating a VAR, which is not feasible in panels where N is not much smaller than T.

Therefore, the SURADF test of Breuer et al (2001, 2002) has been calculated additionally. The first step of the SURADF test, a SUR estimation of the N-equation system (6) without any further restriction, is the same as that for the MADF test. The SURADF, however, proceeds by testing each single estimated ρ_i for stationarity as in a usual univariate ADF test, ie the N null hypotheses H_{0i}: $\rho_i = 0$ are tested separately against the respective alternatives H_{Ai}: $\rho_i < 0$. Thus, the SURADF is in some sense a univariate rather than a panel unit root test. It is, however, applicable only to panels because the panel is necessary for the SUR estimation as a first step. More importantly, the SURE insures that, in contrast to pure univariate unit root tests, information of the whole panel is used. Breuer et al (2002) show that the power of the SURADF test is much higher than that of a simple ADF test if cross-correlations are present, which appears to be the case for the panels examined

[4] This is why the Stata freeware file "madfuller" that uses exclusively the response surface estimation should be avoided with these panel sizes.

here. Compared with tests such as the JLR test, the SURADF test has the additional advantage that it can determine which of the time series in the panel is stationary and which is not.

The critical values for the SURADF test must be simulated and are not the same as univariate ADF critical values. For the Monte Carlo simulations for both the MADF and the SURADF tests, it is first necessary to determine the lag length p_i for each variable. Thus, equation (6) has been estimated separately for each individual under the null which is non-stationarity without drift, ie the intercept and the lagged level have been excluded from (6), $\mu_i = \rho_i = 0$. The largest lag has successively been excluded until $\delta_{i,pi}$ was significant. Since SUR estimations require $N < T$, a maximum lag of three has been chosen, which is still rather large given the annual frequency of the data. Using the determined lag lengths p_i for each of the variables, the system (6) has been re-estimated under the null, ie without intercept and lagged level, and the covariance matrix of the residuals has been calculated.

Given the estimated parameters of the lagged differences and the estimated covariance matrix of the residuals, each series in the panel has been artificially generated with randomly drawn error terms that are normally distributed for each individual using (6) where $\mu_i = \rho_i = 0$. In order to reduce sensitivity to initial conditions, $T + 50$ artificial observations have been generated for each series, the first 50 of which have been discarded. In order to determine the critical values, a SUR estimation of the system (6) including intercept and lagged level has been performed using the artificially generated variables. The relevant test statistics of the MADF and the SURADF tests have been computed. The procedure has been replicated 10,000 times.

4.2 Results

It is usually found in the literature that real exchange rates are stationary when panel unit root tests are applied. In the first step, an examination is made as to whether this result can be replicated with our data. Panels of ten economy-wide real exchange rates are used whose numéraire country is either the USA or Germany. Table 1 shows that the LL, the IPS and the MADF tests reject non-stationarity of real exchange rates for all panels.

By rejecting non-stationarity and thus confirming the validity of PPP, these results are in line with most of those in the literature. The IPS and the MADF, however, are designed to reject the null as soon as there is at least one stationary series in the panel. Their results indicating stationarity might therefore be caused by just one stationary series among N-1 non-stationary ones. In order to examine this, the SURADF test has been performed for these panels. Table 2 does indeed reveal that there is no panel in which the majority of series is found to be stationary. When Germany is chosen as numéraire country, non-stationarity of only one out of ten real exchange rates is rejected, while tests with the USA as numéraire country yield five rejections. Breuer et al (2001, 2002) obtain similar results when applying the SURADF test. The evidence in support of PPP in the literature may thus be overstated.

Testing for stationarity of domestic relative prices amounts to an examination of condition 1. To this end, relative price series of 17 sectors are combined in panels, one for each of the 11 countries. The results, which are presented in Table 3, reconfirm for most countries what the preliminary visual inspection of US relative prices suggested: Regardless of the test statistic (LL, IPS or MADF), the null of non-stationarity cannot be rejected for most of the panels. The implications for condition 1 are aggravated by the fact that a rejection is indeed necessary but not sufficient for condition 1 to be fulfilled: Condition 1 requires that all the relative price series are stationary, and these test results suggest that, for most countries, there is no evidence of even one series being stationary.

The SURADF results presented in Table 4 reveal that there is no panel in which more than three out of 17 series are found to be stationary. It is reconfirmed, moreover, that the majority of panels contain no stationary series at all.[5] Taking into consideration the fact that a 5 % significance level is chosen, with 181 series being subject to the SURADF test, nine rejections would be expected under the assumption of independence even if there were no stationary series at all. In fact, there are only seven rejections. In sum, all the tests provide substantial evidence that relative prices are non-stationary. This implies that condition 1 for the validity of PPP is clearly not fulfilled. Devereux’s (1997) trend movements of relative prices do exist. This could be a consequence of significant technological and/or demand shocks.

PPP would nevertheless hold if condition 2 were fulfilled. This could, for instance, be the case if the weights of the sectors were similar across countries. Condition 2 is examined with panels, each of which comprises 17 series of weighted domestic relative prices of a given country related to the same variable in a numéraire country. As before, the numéraire country is either the USA or Germany. Table 5 shows that non-stationarity for these variables is rejected more often than was the case for pure domestic relative prices in Table 3.

---

[5] Note that the MADF test which can be seen as a pooled equivalent of the SURADF tests failed to find the single stationary series in the Belgian panel, but managed to detect the single stationary relative price among 16 non-stationary ones in the Canadian panel. Similar comparisons can be made throughout the paper.

The outcome of the tests appear, however, to be heavily dependent on the choice of numéraire country (with, generally, more rejections when the USA is chosen as numéraire country) and especially on the choice of test: While the IPS rejects non-stationarity in only one panel, there are six rejections when the LL test is used and as many as 12 when the MADF test is used. Thus, the MADF test – which, contrary to the LL and the IPS test, takes account of cross-correlations – suggests that some of the panels consist exclusively of non-stationary time series while other panels contain stationary time series. In much the same way as condition 1, condition 2 requires, however, that all the series of the panel are stationary. Applying SURADF tests, Table 6 reveals that this is clearly not the case for any panel even if the USA is chosen as numéraire country. Thus, condition 2 is not fulfilled either.

This is still innocuous for PPP if condition 3 is fulfilled. For an examination of condition 3, the time series of a given panel investigated for condition 2, which correspond to the expressions in square brackets in (5), are summed up across sectors. For each country, one time series results. These series are combined in a country panel. Two such panels have been computed, one using the USA as numéraire country, the other one using Germany instead. If these panels were found to be stationary, the (q×1) vector $[1, 1, ... , 1]'$ would constitute a cointegrating vector for the q non-stationary series of the individual sectors, for some reason. The results obtained with the LL, IPS, MADF, and SURADF tests demonstrate, however, that non-stationarity cannot be rejected for any of the series under investigation (see Tables 7 and 8). The evidence thus overwhelmingly suggests that condition 3 is not fulfilled either.[6]

4.3 Robustness checks

In this section, the robustness of the results is checked in two ways: first, by testing another aggregate for stationarity and, second, by repeating the tests of section 4.2 with another numéraire sector. For reasons of space, the results are not presented in detail, although they are, of course, available from the author upon request.

As an intermediate step between testing for condition 1 and for condition 2, one could test for the stationarity of the weighted domestic relative prices in each country, $\alpha_{ik}(p_{ikt} - p_{i0t})$. This is not trivial because the weights have been computed separately for each year as α_{ikt}

[6] These results reconfirm, by the way, that condition 2 (and condition 1) is not fulfilled. Nevertheless, tests of conditions 1 and 2 have been important because, apart from strengthening the evidence, they can be interpreted easily in economic terms and provide an insight into the reasons for the observed non-stationarity of the sum.

so that they may display trends as well. In particular, one could imagine that a rise in the price of sector *k* relative to the price of the numéraire sector may reduce demand for sector *k* goods, which, in turn, reduces the weight of sector *k*. Thus, a trend in relative prices could be offset by an opposite trend in the corresponding weight. It is, however, doubtful whether the substitutibility between sectors is sufficiently large for a complete elimination of trends in relative prices. Furthermore, it has been shown that condition 2 is not fulfilled. This suggests that trends in relative prices may indeed not be eliminated by trends in weights. In order to dispel any doubts, sector panels of weighted domestic relative prices have been formed, one for each country, and they have been subjected to unit root tests. While the LL test rejects non-stationarity for the majority of countries, there is not a single rejection when the more powerful IPS test is used. Using the MADF test yields rejections for four out of 11 countries but the SURADF test shows that, at most, three out of the 17 sectoral time series of a panel are found to be stationary. In sum, there is substantial evidence of weighted domestic relative prices being non-stationary.

As a second robustness check, all the tests of conditions 1-3 have been repeated with a set of relative prices that use the “Machinery and equipment” sector as numéraire sector. This may be advisable for two reasons. First, since the choice of the numéraire country apparently affects the results of stationarity tests of the real exchange rate (cf Coakley/Fuertes, 2000, Papell/Theodoridis, 2001), the choice of the numéraire sector could prove important as well. Second, one may wonder whether the tradability of goods of the numéraire sector affects the results. While the “Finance, insurance, real estate and business services” sector is generally classified as non-tradable, the “Machinery and equipment” sector is usually said to comprise tradable goods.

It proves to be the case, however, that, while the results concerning the stationarity of individual panels or series often depend on the choice of the numéraire sector, the general results concerning conditions 1-3 remain unchanged. Tests of condition 1 again yield very few rejections, especially when the IPS and the SURADF tests are applied. When testing for condition 2, there are distinctively fewer rejections than was the case with “Finance, insurance, real estate and business services” as numéraire sector regardless of the chosen test. Most strikingly, the LL, IPS, MADF and SURADF tests of condition 3 again do not yield a single rejection. In sum, the results clearly reconfirm that none of conditions 1-3 for the validity of PPP is fulfilled.

5. How can the conflicting evidence be reconciled?

The evidence which has been accumulated in this paper uniformly reveals that none of the three conditions for PPP to hold is fulfilled. Most of the components of the real exchange rate are non-stationary and there is no appropriate cointegrating vector for them. Thus, the second-term sum of the right-hand side of equation (5) is found to be non-stationary for all countries without exception. While these pieces of evidence have been consistent among themselves, they conflict with the finding that real exchange rates as a whole usually turn out to be stationary when they are examined with panel unit root tests – a result that has also been confirmed in this paper.

The conundrum could be resolved if the LOP did not hold. Then, the first term on the right-hand side of (5), r_{i0t}, would be non-stationary and could be cointegrated with the second-term sum with the cointegrating vector [1, 1]′. A stationary real exchange rate would result. Such a solution, however, would raise new questions because it implies, first, that PPP holds although, on the one hand, the LOP does not and, on the other, relative price trends exist. There is no economic rationale for such a case. Second, there is no economic reason either for the two terms of the right-hand side to be cointegrated with vector [1, 1]′.

Accordingly, it is not possible to find much evidence in favour of this hypothesis: Sector-specific real exchange rates have been calculated using the price indices of the numéraire sector and the IFS nominal exchange rate data. The series have been combined in country panels. Results of unit root tests are presented in Tables 9 and 10. With one exception, the LL, IPS and MADF tests reject non-stationarity, which supports the LOP but is inconsistent with the cointegration hypothesis.[7] These results are qualified by the SURADF test, which suggests that only two of the series in the panels are stationary. This, however, does not support the cointegration hypothesis either because, using the SURADF test, most real exchange rates have been found to be stationary as well (cf Table 2).

A second, more plausible, hypothesis for resolving the conflict is that changes in the first term of (5) are large in relation to the changes in the second term and would therefore mask any non-stationarity that comes from the second term. The variance in the difference of the first-term series is indeed 10 times to 100 times as large as the variance in the difference of the second term if the USA is chosen as numéraire country, and it is still 1.3 times to 50 times as large when Germany is the numéraire country. This result clearly reflects the

[7] Note the implication that the LOP holds for some of the series at least although the numéraire sector, “Finance, insurance, real estate and business services”, is usually classified as non-tradable.

considerable movements of nominal exchange rates which are part of the first term but not of the second.

The evidence is consistent with the ideas put forward by Bayoumi/MacDonald (1999) in a related context. These suggest that the observed mean-reversion in real exchange rates is generated by stationary monetary shocks. Such monetary shocks, however, mask long-run trends in the real exchange rate which are created by real factors. In the context of equation (5), the first term on the right-hand side will be driven by the monetary shocks because it represents a relative price between the same good in two different countries with possibly two different monetary policies, while the second-term sum could be driven by real factors that change relative prices between different goods. From this perspective, equation (5) can be seen as one possible way of following Bayoumi/MacDonald's (1999) concluding call to identify and quantify the effects of these two shocks. Bayoumi/MacDonald's (1999) hypothesis implies that PPP does not hold, but that the non-stationary component in real exchange rates is so small that unit root tests cannot detect it.

There is a third way of resolving the conundrum. If the LOP holds, the first term on the right-hand side of (5) will be stationary while it has been shown that the second-term sum is non-stationary. Thus, the real exchange rate consists of a stationary and a non-stationary component. It is well-known, however, that unit root tests are heavily biased in favour of rejecting non-stationarity when they are applied to variables with such a two-component character; see, for example, Blough (1992). The reason for the size bias is that these variables can be represented as variables with negative MA(1) errors; see, for example, Clark (1988). For the specific case of the LL and the IPS test, Im et al (2003) show, in their Table 6, that even a small negative MA(1) parameter can cause considerable size biases in both tests if too small a lag length is chosen. Results of unit root tests that suggest stationarity of the real exchange rate may therefore be due to a size bias caused by the possible two-component character of the real exchange rate.

This is a hypothesis that has already been suggested by Engel (2000). He found that the Balassa-Samuelson hypothesis implies a similar two-component structure for the real exchange rate, one component being mean-reverting and the other one being non-stationary. Using Monte Carlo simulations he determined that the implied size biases for some univariate unit root tests are sizeable. It is interesting that the same two-component structure of the real exchange rate follows here (if the LOP holds) in a quite different modelling environment. While the Balassa-Samuelson hypothesis is based on the existence of non-tradables, the two-component structure of the real exchange rate results here even if the LOP holds for every single good.

6. Conclusions

The paper examines an old but rarely mentioned precondition for purchasing power parity to hold. As Devereux (1997) puts it, if "... the composition of price indices differs across countries, ... trend movements in relative goods prices will lead to persistent deviations from PPP". Until now, there does not appear to have been a systematic empirical examination of whether or not this is the case. In order to formalize the argument, economy-wide price indices have been disaggregated into price indices for single goods. It becomes apparent that the real exchange rate consists of two components; first, the real exchange rate for a single good and, second, a weighted sum of relative prices between *different* goods. PPP requires that the real exchange rate as a whole is stationary. Since stationarity of the first component, the real exchange rate of a single good, implies that the law of one price holds, the second component should be stationary as well in order to maintain the validity of PPP.

Using a battery of panel unit root tests, it is first shown that relative prices between different goods are predominantly non-stationary. Devereux's (1997) trend movements in relative goods prices do exist. In a second step, it is shown that this does not change if weighted relative prices are related to their foreign counterparts. This could be a consequence of differences in the composition of price indices across countries. In a third step, it is found that the non-stationarity does not aggregate out. The second component as a whole is non-stationary.

The results imply that none of three conditions for PPP to hold is fulfilled. While the results are consistent among themselves, they are at odds with the finding of many papers that panel unit root tests suggest the stationarity of real exchange rate as a whole – a result that has been confirmed here. Two ways of solving the conundrum are proposed. Since it is found that the (variance in the difference of) the non-stationary second component is small compared with the possibly stationary first component, the latter could simply mask the former. This has been suggested by Bayoumi/MacDonald (1999) in a related context. Alternatively, the fact that panel unit root tests find that real exchange rates are mean-reverting could be due to a bias. This bias arises if variables are tested that consist of one stationary and one non-stationary component, which is exactly what this paper suggests if the LOP holds. In this respect, the paper supports Engel's (2000) suggestion that unit root tests are subject to such a size bias when they are applied to real exchange rates. In contrast to Engel (2000), who derives this result in a Balassa-Samuelson framework, however, it is found here that this follows even if the LOP holds for every single good.

Appendix: Some additional information on the data

The following sectoral subdivision of the economy has been chosen.

No.	Sector
1	Agriculture, hunting, forestry and fishing
2	Mining and quarrying
3	Food products, beverages and tobacco
4	Textiles, textile products, leather and footwear
5	Wood and products of wood and cork
6	Pulp, paper, paper products, printing and publishing
7	Chemical, rubber, plastics and fuel products
8	Other non-metallic mineral products
9	Basic metals and fabricated metal products
10	Machinery and equipment
11	Transport equipment
12	Manufacturing nec; recycling
13	Electricity, gas and water supply
14	Construction
15	Wholesale and retail trade; restaurants and hotels
16	Transport and storage and communication
17	Community social and personal services
18	Finance, insurance, real estate and business services

Of the two sets of data for Japan provided in the STAN dataset, the one which included more recent data has been used. For Germany, data up to and including 1991 is given for Western Germany, and data from 1991 onwards for Germany as a whole. Price indices up to and from 1991 have been linked.

References

Balassa, B, 1964, The purchasing-power parity doctrine: a reappraisal, *The Journal of Political Economy*, Vol 72, 584-596.

Bayoumi, T, R MacDonald, 1999, Deviations of exchange rates from purchasing power parity: a story featuring two monetary unions, *IMF Staff Papers*, Vol 46, 89-102.

Blough, S R, 1992, The relationship between power and level for generic unit root tests in finite samples, *Journal of Applied Econometrics*, Vol 7, 295-308.

Breuer, J B, 1994, An assessment of the evidence on purchasing power parity, in J Williamson (ed), *Estimating Equilibrium Exchange Rates*, Institute for International Economics, 245-277.

Breuer, J B, R McNown, M S Wallace, 2001, Misleading inferences from panel unit-root tests with an illustration from purchasing power parity, *Review of International Economics*, Vol 9, 482-493.

Breuer, J B, R McNown, M Wallace, 2002, Series-specific unit root tests with panel data, *Oxford Bulletin of Economics and Statistics*, Vol 64, 527-546.

Breusch, T, A Pagan, 1980, The LM test and its applications to model specification in econometrics, *Review of Economic Studies*, Vol 47, 239-254.

Campbell, J Y, P Perron, 1991, Pitfalls and opportunities: what macroeconomists should know about unit roots, *NBER Macroeconomics Annual 1991*, 141-201.

Canzoneri, M B, R E Cumby, B Diba, 1999, Relative labour productivity and the real exchange rate in the long run: evidence for a panel of OECD countries, *Journal of International Economics*, Vol 47, 245-266.

Choi, I, 2001, Unit root tests for panel data, *Journal of International Money and Finance*, Vol 20, 249-272.

Clark, P K, 1988, Nearly redundant parameters and measures of persistence in economic time series, *Journal of Economic Dynamics and Control*, Vol 12, 447- 461.

Coakley, J, A-M Fuertes, 2000, Is there a base currency effect in long-run PPP?, *International Journal of Finance and Economics*, Vol 5, 253-263.

De Gregorio, J, A Giovannini, H C Wolf, 1994, International evidence on tradables and nontradables inflation, *European Economic Review*, Vol 38, 1225-1244.

DeLoach, S B, 2001, More evidence in favor of the Balassa-Samuelson hypothesis, *Review of International Economics*, Vol 9, 336-342.

Devereux, M B, 1997, Real exchange rates and macroeconomics: evidence and theory, *Canadian Journal of Economics*, Vol 30, 773-808.

Dickey, D A, W A Fuller, 1981, Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, Vol 49, 1057-1072.

Dornbusch, R, 1987, Purchasing power parity, in J Eatwell, M Milgate, P Newman (eds), *The New Palgrave: A Dictionary of Economics*, Macmillan Press, 1075-1085.

Engel, C, 1999, Accounting for U. S. real exchange rates, *Journal of Political Economiy*, Vol 107, 507-538.

Engel, C, 2000, Long-run PPP may not hold after all, *Journal of International Economics*, Vol 57, 243-273.

Fleissig, A R, J Strauss, 2000, Panel unit root tests of purchasing power parity for price indices, *Journal of International Money and Finance*, Vol 19, 489-506.

Froot, K A, K Rogoff, 1995, Perspectives on PPP and long-run real exchange rates, in G M Grossman, K Rogoff (eds), *Handbook of International Economics*, Vol 3, Elsevier, 1647-1688.

Higgins, M, E Zakrajšek, 1999, Purchasing power parity: three stakes through the heart of the unit root null, Federal Reserve Bank of New York Staff Reports, No. 80.

Ho, T-W, 2002, Searching stationarity in the real exchange rates: application of the SUR estimator, *Open Economies Review*, Vol 13, 275-289.

Hsieh, D A, 1982, The determination of the real exchange rate: The productivity approach, *Journal of International Economics*, Vol 12, 355-362.

Im, K S, M H Pesaran, Y Shin, 2003, Testing for unit roots in heterogeneous panels, *Journal of Econometrics*, Vol 115, 53-74.

Kuo, B-S, A Mikkola, 2001, How sure are we about purchasing power parity? Panel evidence with the null of stationary real exchange rates, *Journal of Money, Credit, and Banking*, Vol 33, 767-789.

Lan, Y, 2002, The explosion of purchasing power parity, in M Manzur (ed), *Exchange Rates, Interest Rates and Commodity Prices*, Edward Elgar, 9-38.

Levin, A, C-F Lin, C-S J Chu, 2002, Unit root tests in panel data: asymptotic and finite-sample properties, *Journal of Econometrics*, Vol 108, 1-24.

Obstfeld, M, K Rogoff, 1996, *Foundations of International Economics*, MIT Press.

O'Connell, P G J, 1998, The overvaluation of purchasing power parity, *Journal of International Economics*, Vol 44, 1-19.

Papell, D H, H Theodoridis, 2001, The choice of numeraire currency in panel tests of purchasing power parity, *Journal of Money, Credit, and Banking*, Vol 33, 790-803.

Parsley, D C, H A Popper, 2001, Official exchange rate arrangements and real exchange rate behavior, *Journal of Money, Credit, and Banking*, Vol 33, 976-993.

Rogoff, K, 1996, The purchasing power parity puzzle, *Journal of Economic Literature*, Vol 34, 647-668.

Samuelson, P A, 1964, Theoretical notes on trade problems, *The Review of Economics and Statistics*, Vol 46, 145-154.

Sarno, L, M P Taylor, 1998, Real exchange rates under the recent float: unequivocal evidence of mean reversion, *Economics Letters*, Vol 60, 131-137.

Sarno, L, M P Taylor, 2002, *The Economics of Exchange Rates*, Cambridge University Press.

Sarno, L, M P Taylor, I Chowdhury, 2004, Nonlinear dynamics in deviations from the law of one price: a broad-based empirical study, *Journal of International Money and Finance*, Vol. 23, 1-25.

Strauss, J, T Yigit, 2003, Shortfalls of panel unit root testing, *Economics Letters*, Vol 81, 309-313.

Taylor, M P, L Sarno, 1998, The behaviour of real exchange rates during the post-Bretton Woods period, *Journal of International Economics*, Vol 46, 281-312.

Wei, S-J, D C Parsley, 1995, Purchasing power disparity during the floating rate period: exchange rate volatility, trade barriers and other culprits, NBER Working Paper No. 5032.

Wu, J-L, S Wu, 2001, Is purchasing power parity overvalued?, *Journal of Money, Credit, and Banking*, Vol 33, 804-812.

Table 1. Panel unit root tests of real exchange rates, country panels

Numéraire country	LL	IPS	MADF
USA	-2.99*	-2.68*	111.62*
Germany	-2.15*	-1.93*	102.30*

The t-star statistic of the LL test and the Psi(t-bar)-statistic of the IPS test are shown. A star indicates a rejection at the 5% significance level.

Table 2. SURADF tests of real exchange rates

	Numéraire country: USA	Numéraire country: Germany
Austria	-7.16*	-5.56*
Belgium	-9.49*	-3.02
Canada	-2.86	-2.86
Denmark	-7.39*	-2.16
Finland	-6.59*	-2.27
Germany	-7.63*	-
Italy	-5.87	-3.03
Japan	-3.93	-2.37
Korea	-2.35	-4.27
UK	-2.77	-3.54
USA	-	-3.05

A star indicates a rejection at the 5% significance level. Note that, owing to differing correlation structures in the two panels vis-á-vis third countries, the result for Germany when the USA is numéraire country may differ from the result for the USA when Germany is numéraire country.

Table 3. Panel unit root tests of domestic relative prices, sector panel

	LL	IPS	MADF
Austria	-1.49	0.50	396.42
Belgium	-0.40	1.99	403.21
Canada	-1.00	1.09	564.98*
Denmark	0.62	0.78	399.20
Finland	-1.43	-0.80	125.25
Germany	2.44	3.40	245.30
Italy	-4.79*	-2.93*	262.16
Japan	-2.47*	-0.45	215.61
Korea	0.50	2.34	449.45*
UK	0.52	1.74	255.75
USA	2.42	2.34	255.03

Numéraire sector: “Finance, insurance, real estate and business services”. The t-star statistic of the LL test and the Psi(t-bar)-statistic of the IPS test are shown. A star indicates a rejection at the 5% significance level.

Table 4. SURADF tests of domestic relative prices

Sector No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Austria	-1.22	-2.56	-2.46	-3.86	-5.49	-3.80	0.64	-5.58	-2.57	-3.91	-5.99	-2.26	1.73	-5.69	-5.45	-1.68	-5.93
Belgium	0.02	-2.77	-3.51	-0.22	-2.61	-5.29	-6.84	-5.90	-7.85	-8.98*	-7.29	-1.67	1.71	-4.07	0.94	-4.96	-4.50
Canada	-6.34	-3.43	-6.79	-5.95	-3.93	-9.03*	-4.91	-2.82	-6.00	-6.63	-4.92	-7.71	-6.49	-4.22	-4.80	-1.32	-3.31
Denmark	-1.84	-2.61	-0.79	-1.39	-7.06	0.03	2.17	-6.85	-6.00	-2.83	-3.86	-7.02	-7.41	-3.91	-3.97	1.05	-4.40
Finland	-1.26	-0.09	1.96	-3.12	-0.30	-5.28	-2.25	-2.62	-1.91	-0.97	-2.13	-1.52	-1.44	-6.28	-1.35	-2.74	-2.60
Germany	-2.88	1.09	-9.42*	-4.75	-5.37	-7.53	-3.68	-5.99	-4.55	-5.41	-7.25	2.63	-2.59	-6.59	-9.66*	0.42	-12.45*
Italy	-0.73	-2.37	0.27	-0.02	-1.86	-0.92	0.21	-2.43	-1.88	-0.92	-5.25	-0.98	-6.34	2.95	1.60	0.10	-9.81*
Japan	-6.38	-5.61	-6.88	-2.78	-5.50	-4.01	-2.00	-2.15	-3.50	-1.03	-2.11	-4.66	-6.66	-1.98	-5.78	1.76	-1.54
Korea	-0.26	-1.43	-3.60	-1.02	-3.42	-2.95	-2.02	0.40	-0.50	1.25	-0.56	-3.58	-1.14	-8.84*	1.15	1.25	-5.89
UK	-1.17	-1.18	-3.13	-5.83	-3.05	-2.67	-4.11	-2.18	-4.39	-2.50	-2.13	-2.87	-0.10	-3.57	-6.71	4.96	-5.93
USA	-3.01	-0.68	-4.86	-6.21	-5.43	-4.86	-4.21	-3.98	-2.26	2.70	-4.56	-3.02	-2.74	-2.59	0.39	-1.69	-1.84

A key to the sectors is given in the appendix. Numéraire sector: “Finance, insurance, real estate and business services”. A star indicates a rejection at the 5% significance level.

Table 5. Panel unit root tests of weighted domestic relative prices in relation to weighted relative prices in a numéraire country, sector panel

	Numéraire country: USA			Numéraire country: Germany		
	LL	IPS	MADF	LL	IPS	MADF
Austria	1.31	1.94	394.28*	-1.06	1.12	601.71*
Belgium	0.99	3.98	249.88	-5.76*	-1.61	1692.62*
Canada	-0.06	0.42	408.36*	-1.18	0.24	284.07
Denmark	0.90	0.99	851.40*	-0.77	1.52	193.76
Finland	-2.10*	-0.23	198.93	0.13	0.84	294.92
Germany	2.98	5.00	445.96*	-	-	-
Italy	-3.70*	-0.36	868.30*	-2.12*	0.12	426.53*
Japan	-2.90*	-0.33	457.11*	-1.36	-0.01	182.45
Korea	-2.08*	-1.70*	536.20*	-0.63	0.89	410.94
UK	0.00	1.54	775.50*	-0.65	-0.14	359.69
USA	-	-	-	2.98	5.00	445.96*

Numéraire sector: "Finance, insurance, real estate and business services". The t-star statistic of the LL test and the Psi(t-bar)-statistic of the IPS test are shown. A star indicates a rejection at the 5% significance level.

Table 6. SURADF tests of weighted domestic relative prices in relation to weighted relative prices in a numéraire country

Sector No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Numéraire country: USA																	
Austria	-1.36	-4.32	-0.89	-2.65	-1.35	-0.81	-8.55*	-5.72	-2.40	0.72	-6.04	-9.89*	-2.99	-4.96	-1.70	-4.92	-3.36
Belgium	-5.95	-5.50	-6.25	-4.68	-2.84	-3.42	-3.46	-7.60	-6.52	1.71	-6.64	-6.44	-2.46	-4.07	-1.26	-0.42	-5.16
Canada	-5.47	-5.76	-6.34	-8.16*	-1.68	-5.49	-5.07	-5.87	-5.04	3.77	-1.82	-6.82	-0.99	-4.29	-1.50	-4.77	-4.71
Denmark	-1.33	-2.11	-0.69	-8.31	-8.59*	-6.35	-6.33	-8.48	-3.81	1.92	-4.06	-1.76	-2.69	-2.43	-1.27	-6.89	-3.52
Finland	-0.86	-3.90	2.30	-7.19	-0.79	-2.72	-3.09	-2.51	-4.70	-2.20	-5.08	-2.00	-5.42	-5.63	-2.90	-3.80	-3.60
Germany	-11.02*	-4.79	-4.49	-2.01	-4.95	-2.07	-3.56	-2.38	-4.82	-0.65	-6.92	-4.92	-5.73	-2.10	0.98	-3.09	-7.50
Italy	-1.49	-5.88	-1.18	-1.58	-2.36	1.48	-7.94*	-1.83	-1.58	2.03	-7.89	-1.63	0.10	1.68	-11.70*	-5.27	-7.93*
Japan	-4.16	-3.77	-8.18*	-6.30	-3.54	-5.08	-9.69*	-7.71*	-6.94	-9.77*	-6.05	-6.49	-3.80	-2.36	-1.13	-3.83	-2.63
Korea	-3.54	-4.10	-2.66	-6.27	-11.53*	-1.23	-4.88	-1.72	-10.17*	-0.69	-1.96	-6.29	-4.82	-3.23	-7.78	-3.96	-16.87*
UK	-12.78*	-3.08	-1.14	-8.39	-2.57	-1.76	-1.75	-7.35	-8.92*	0.05	-3.16	-1.98	-3.91	-2.14	-1.82	0.64	-4.31
Numéraire country: Germany																	
Austria	-1.45	-1.22	-3.00	-4.13	-7.45	-2.91	-7.29	-8.11	-9.62*	-7.22	-8.89*	-2.84	-3.09	-2.62	-0.85	-4.50	-5.87
Belgium	-0.96	-1.85	-3.71	0.47	-2.86	-1.81	-8.95*	-14.66*	-10.30*	-4.85	-5.22	2.31	-0.70	-2.24	-2.89	4.15	-6.28
Canada	-9.04*	-1.19	-9.12*	-7.50	-4.67	-2.66	-4.04	-6.51	-3.06	-7.02	-5.15	-0.58	-1.75	-3.61	1.37	-2.42	-6.78
Denmark	-1.07	-6.17	-2.46	-5.37	-5.55	-3.47	-3.05	-3.79	-0.79	-5.35	-6.93	-3.01	-3.23	-2.51	-2.64	-4.89	-0.93
Finland	-1.58	-10.27*	1.10	-4.14	-6.98	-3.61	-3.51	-0.83	-1.72	-1.15	-2.53	-1.92	-2.11	-3.74	-1.62	-4.70	-6.60
Italy	-3.27	-5.48	-0.22	0.99	-2.75	0.14	-7.39	-2.72	-3.96	-2.25	-2.71	-2.19	-2.20	-0.80	1.44	-5.76	-5.68
Japan	-2.00	-1.63	-5.17	-0.33	-5.43	-6.11	-3.54	-5.85	-6.93	-0.58	-3.96	-0.06	-5.14	-2.72	0.26	-1.27	-2.57
Korea	-4.06	-5.67	-4.01	-4.10	-6.08	-2.90	-3.68	-1.19	-2.53	2.47	-1.16	-1.61	-2.70	-4.12	1.79	-6.21	-14.96*
UK	-4.36	-3.67	-2.85	-6.71	-2.46	-4.72	-8.01	-12.62*	-8.49*	-5.56	-1.20	-9.64*	-4.61	-3.99	-4.37	0.09	-4.49
USA	-11.02*	-4.79	-4.49	-2.01	-4.95	-2.07	-3.56	-2.38	-4.82	-0.65	-6.92	-4.92	-5.73	-2.10	0.98	-3.09	-7.50

A key to the sectors is given in the appendix. Numéraire sector: “Finance, insurance, real estate and business services”. A star indicates a rejection at the 5% significance level.

Table 7. Panel unit root tests of the second-term sum in (5), country panels

Numéraire country	LL	IPS	MADF
USA	-0.28	1.51	39.42
Germany	0.36	1.54	62.37

Numéraire sector: "Finance, insurance, real estate and business services". The t-star statistic of the LL test and the Psi(t-bar)-statistic of the IPS test are shown. A star indicates a rejection at the 5% significance level.

Table 8. SURADF tests of the second-term sum in (5)

	Numéraire country: USA	Numéraire country: Germany
Austria	-3.37	-3.42
Belgium	-2.03	-4.08
Canada	-1.18	-4.30
Denmark	-3.48	-2.21
Finland	-2.32	0.21
Germany	-2.61	-
Italy	-2.29	-0.23
Japan	-3.17	-1.92
Korea	-3.85	-0.59
UK	-1.37	-3.02
USA	-	-5.02

Numéraire sector: "Finance, insurance, real estate and business services". A star indicates a rejection at the 5% significance level. Note that, owing to differing correlation structures in the two panels vis-á-vis third countries, the result for Germany when the USA is numéraire country may differ from the result for the USA when Germany is numéraire country.

Table 9. Panel unit root tests of the sector-specific real exchange rate of the numéraire sector, country panels

Numéraire country	LL	IPS	MADF
USA	-4.29*	-3.22*	100.65*
Germany	-3.15*	-2.40*	70.65

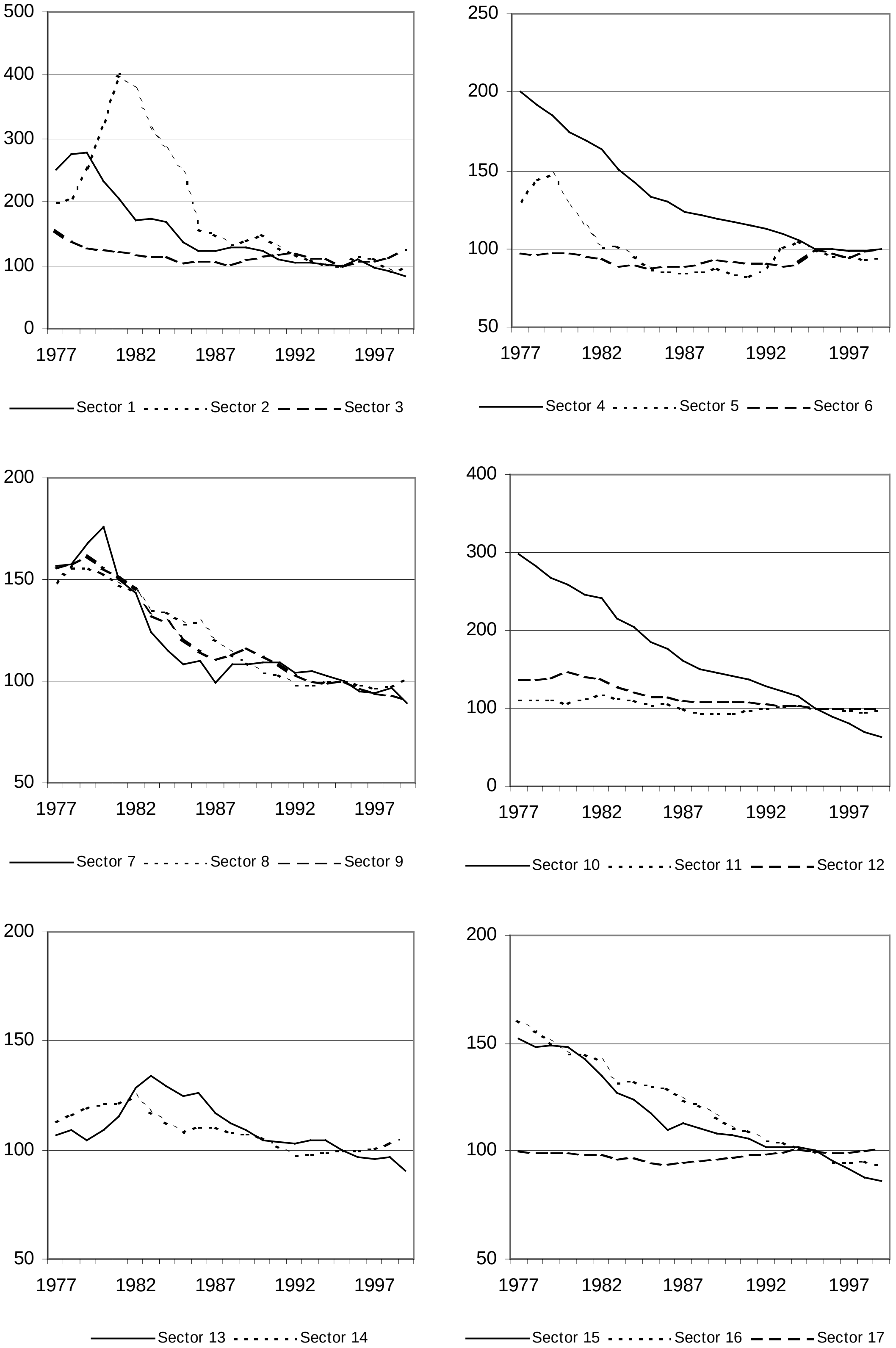
Numéraire sector: “Finance, insurance, real estate and business services”. The t-star statistic of the LL test and the Psi(t-bar)-statistic of the IPS test are shown. A star indicates a rejection at the 5% significance level.

Table 10. SURADF tests of the sector-specific real exchange rate of the numéraire sector

	Numéraire country: USA	Numéraire country: Germany
Austria	-5.94	-2.19
Belgium	-7.56*	-3.68
Canada	-2.37	-3.55
Denmark	-8.16*	-1.59
Finland	-5.89	-3.66
Germany	-5.96	-
Italy	-5.11	-2.86
Japan	-5.04	-4.06
Korea	-2.63	-3.15
UK	-2.88	-2.67
USA	-	-4.12

Numéraire sector: “Finance, insurance, real estate and business services”. A star indicates a rejection at the 5% significance level. Note that, owing to differing correlation structures in the two panels vis-á-vis third countries, the result for Germany when the USA is numéraire country may differ from the result for the USA when Germany is numéraire country.

Figure 1. Sectoral price indices relative to the price index of the "Finance, insurance, real estate and business services" sector in the USA, 1995 = 100



A key to the sectors is given in the appendix.

The following Discussion Papers have been published since 2003:

Series 1: Studies of the Economic Research Centre

January	2003	Testing mean-variance efficiency in CAPM with possibly non-gaussian errors: an exact simulation-based approach	Marie-Claude Beaul Jean-Marie Dufour Lynda Khalaf
January	2003	Finite-sample distributions of self-normalized sums	Jeong-Ryeol Kim
January	2003	The stock return-inflation puzzle and the asymmetric causality in stock returns, inflation and real activity	Jeong-Ryeol Kim
February	2003	Multiple equilibrium overnight rates in a dynamic interbank market game	Jens Tapking
February	2003	A comparison of dynamic panel data estimators: Monte Carlo evidence and an application to the investment function	Andreas Behr
March	2003	A Vectorautoregressive Investment Model (VIM) And Monetary Policy Transmission: Panel Evidence From German Firms	Joerg Breitung Robert S. Chirinko Ulf von Kalckreuth
March	2003	The international integration of money markets in the central and east European accession countries: deviations from covered interest parity, capital controls and inefficien-Cies in the financial sector	Sabine Herrmann Axel Jochem
March	2003	The international integration of foreign exchange markets in the central and east European accession countries: speculative efficiency, transaction costs and exchange rate premiums	Sabine Herrmann Axel Jochem

March	2003	Determinants of German FDI: New Evidence from Micro-Data	Claudia Buch Jörn Kleinert Farid Toubal
March	2003	On the Stability of Different Financial Systems	Falko Fecht
April	2003	Determinants of German Foreign Direct Investment in Latin American and Asian Emerging Markets in the 1990s	Torsten Wezel
June	2003	Active monetary policy, passive fiscal policy and the value of public debt: some further monetarist arithmetic	Leopold von Thadden
June	2003	Bidder Behavior in Repo Auctions without Minimum Bid Rate: Evidence from the Bundesbank	Tobias Linzert Dieter Nautz Jörg Breitung
June	2003	Did the Bundesbank React to Stock Price Movements?	Martin T. Bohl Pierre L. Siklos Thomas Werner
15	2003	Money in a New-Keynesian model estimated with German data	Jana Kremer Giovanni Lombardo Thomas Werner
16	2003	Exact tests and confidence sets for the tail coefficient of α-stable distributions	Jean-Marie Dufour Jeong-Ryeol Kurz-Kim
17	2003	The Forecasting Performance of German Stock Option Densities	B R Craig, E Glatzer, J Keller, M Scheicher
18	2003	How wacky is the DAX? The changing structure of German stock market volatility	Jelena Stapf Thomas Werner

1	2004	Foreign Bank Entry into Emerging Economies: An Empirical Assessment of the Determinants and Risks Predicated on German FDI Data	Torsten Wezel
2	2004	Does Co-Financing by Multilateral Development Banks Increase "Risky" Direct Investment in Emerging Markets? – Evidence for German Banking FDI	Torsten Wezel
3	2004	Policy Instrument Choice and Non-Coordinated Monetary Policy in Interdependent Economies	Giovanni Lombardo Alan Sutherland
4	2004	Inflation Targeting Rules and Welfare in an Asymmetric Currency Area	Giovanni Lombardo
5	2004	FDI versus cross-border financial services: The globalisation of German banks	Claudia M. Buch Alexander Lipponer
6	2004	Clustering or competition? The foreign investment behaviour of German banks	Claudia M. Buch Alexander Lipponer
7	2004	PPP: a Disaggregated View	Christoph Fischer

Series 2: Banking and Financial Supervision

1	2003	Measuring the Discriminative Power of Rating Systems	B. Engelmann, E. Hayden, D. Tasche
2	2003	Credit Risk Factor Modeling and the Basel II IRB Approach	A. Hamerle, T. Liebig, D. Rösch
1	2004	Forecasting Credit Portfolio Risk	A. Hamerle, T. Liebig, H. Scheule

Visiting researcher at the Deutsche Bundesbank

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a Ph D and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank
Personalabteilung
Wilhelm-Epstein-Str. 14

D - 60431 Frankfurt
GERMANY