International Network for Economic Research

Working Paper 2010.3

Equilibrium Real Effective Exchange Rates and Real Exchange Rate Misalignments: Time Series vs. Panel Estimates

by

Oliver Hossfeld

(Sparkassen-Finanzgruppe-Chair of Macroeconomics, HHL – Leipzig Graduate School of Management)

Equilibrium Real Effective Exchange Rates and Real Exchange Rate Misalignments: Time Series vs. Panel Estimates

Oliver Hossfeld*

September 2010

Abstract

In this paper we apply the Behavioral Equilibrium Exchange Rate (BEER) approach developed by Clark and MacDonald (1998) to derive equilibrium real effective exchange rates and currency misalignments for the US and its major 16 trading partners over a sample from 1986Q1 to 2006Q4. Cointegration and panel cointegration techniques are applied to derive fully country-specific measures of misalignment and country-specific measures based on panel estimates. We do not only apply the popular first generation panel unit root and panel cointegration tests, but also two recently introduced classes of tests of the second generation: the CIPS (cross-sectionally augmented IPS) panel unit root tests by Pesaran (2007) as well as the error-correction-based tests for panel cointegration by Westerlund (2007), which both account for possible cross-sectional dependencies among the units included in the panel. Using the estimates obtained over a restricted sample, forecasting tests are conducted to assess the relative forecasting performance of pooled vs. heterogeneous estimators. We find that pooling the data delivers more reliable results when calculating equilibrium exchange rates though the implicit homogeneity restriction is statistically rejected. This result is especially remarkable, since we have given the heterogeneous estimator an 'unfair' advantage by choosing the country-specific model (of up to 21 possible ones) with the best out-of-sample performance prior to comparing it to two final panel

*Address: Sparkassen-Finanzgruppe-Chair of Macroeconomics, HHL - Leipzig Graduate School of Management, Jahnallee 59, 04109 Leipzig, Germany. Email: oliver.hossfeld@hhl.de. I thank conference participants in Berlin and Magdeburg for helpful comments and discussions.

specifications. Based on this result, we re-estimate our preferred specification and calculate real effective exchange rate misalignments over the full sample. While we find strong evidence in favor of the Balassa-Samuelson-effect, evidence in favor of other commonly hypothesized fundamentals is weak.

JEL Classification Numbers: C22, C23, F31, F37

1 Introduction

Since its peak value in April 2002 the real value of the US-Dollar has declined steadily, reaching an overall decline of more than 20 percent in early 2008. The question naturally arises whether this drop of the US Dollar and the simultaneous rise and fall of other currencies are in line with underlying fundamentals (if there are any) and so represents movements of the equilibrium exchange rates, or whether they are caused by other factors such as currency speculation. The enormous increase of the US net foreign liabilities, turning the US from the world's largest creditor into the world's largest debtor, raised the question whether the real value of the US Dollar dropped in order to correct this global imbalance and if so, how much more there is to come. Net foreign assets are however just one of the commonly hypothesized fundamentals exerting an influence on the real (effective) exchange rate.

Until the end of the 1990s authors focused on time series analysis to estimate equilibrium real effective exchange rates, assess the influence of certain hypothesized fundamentals, and derive currency misalignments (for instance Faruqee, 1995, Clark and MacDonald, 1998, and MacDonald, 1998). More recently, several authors have taken a panel perspective (for instance Kim and Korhonen, 2005, Villavicencio, 2006, Bénassy-Quéré et al., 2008, and Ricci et al., 2008). Results from MacDonald and Dias (2007) however indicate considerable differences among individual countries' parameter estimates, as well as between individual countries' parameter estimates and their panel equivalents.[1] This may either be caused by true parameter heterogeneity among the countries, by imprecise estimates due to the low number of observations in the time series dimension, by a biased pooled estimator in the case of true parameter heterogeneity, or most probably a combination of them.[2] So, should we put more trust in the individual time series estimates or in the pooled estimates when calculating equilibrium exchange rates and deriving currency misalignments? The answer to this question bears important policy implications. If a country is joining a currency union, the conversion rate of its currency against the one of the monetary union needs to be locked at some value. Unreliable estimates of the proper rate, and, as a consequence, a pos-

[1]Their estimated coefficients for the trade balance to GDP ratio vary between -177.77 for Japan and -1.34 for the US.

[2]In the case of true parameter heterogeneity across countries, country-specific (i.e. heterogeneous) time series estimations are theoretically the first choice, because the pooled estimator is biased if the implicit (parameter) homogeneity restriction is rejected (Maddala et al., 1997). However, if the time series are too short, estimates can be highly inaccurate and estimated coefficients be even incorrectly signed (Baltagi et al., 2008 and Maddala et al., 1997).

sibly severe currency misalignment, could unnecessarily give rise to inflation or deflation in the accession country and even slow down economic growth in the case of prolonged misalignments (Edwards and Savastano, 1999). The 'fundamentally justified' value of a currency is not only of interest to central bankers and government agencies, but also to the private sector (Edwards and Savastano, 1999), because the current deviation of the observed real effective exchange rate from its predicted long-run value may be helpful to predict future exchange rate movements.

Rapach and Wohar (2004) analyze whether the monetary exchange rate model performs better in forecasting bilateral *nominal* exchange rates if the underlying parameter estimates are obtained from panel estimations as compared to country-by-country estimations. They find the pooled estimator to deliver a better forecasting performance, although the underlying poolability hypothesis is statistically rejected. With regard to the empirical modeling of *real* equilibrium exchange rates, Égert (2004) notes that

> 'Across different papers, the whole gamut of fundamentals is used, and, as a corollary, the outcome is sensitive to which particular fundamentals are included in the estimated model. The use of different fundamentals may be a result of different theoretical frameworks or may simply reflect ad hoc choices.'

So, whereas the choice of fundamentals is predetermined by the choice of the monetary exchange rate model in Rapach and Wohar (2004), there is at least some degree of arbitrariness with respect to the selection of determinants in the class of models we consider.[3] Notwithstanding the arbitrariness observed in the relevant literature, it may well be that the importance of certain determinants in explaining real effective exchange rate movements differs across countries.[4]

To assess the reliability of time series versus panel approaches in estimating equilibrium exchange rates we therefore estimate up to 21 specifications for each of the included countries (depending on the univariate time properties of the considered determinants) and twelve panel specifications over a restricted sample from 1986Q1 to 2002Q4, and then compare the relative

[3]Commonly chosen determinants include net foreign assets, measures of relative productivity differentials, relative terms of trade, measures of the openness of a country, the relative stock of government debt and real interest rate differentials. Less frequently, also the real oil price (especially for emerging market countries) has been considered (Amano and van Norden 1998).

[4]However, it may be questionable whether one should call these country-specific determinants '*fundamental*' determinants of real exchange rates.

accuracy of *conditional* forecasts obtained from these models over a hold-back period from 2003Q1 to 2006Q4.[5] In all steps of the analysis we put ourselves once in the position of a time series econometrician who disregards *any* information about the other cross units, and once in the position of a panel econometrician who takes *all* available information into account.

In our analysis we will focus on two 'extreme' estimators: The fully country-specific (i.e. heterogeneous) dynamic OLS (DOLS) estimator and the pooled DOLS estimator.[6] Because we impose homogeneity of the cointegration slopes in the pooled estimations, we formally check whether this restriction holds. However, even a rejection of the poolability null would not necessarily mean that the obtained estimates are worthless. It may well be that the 'benefits of higher estimation precision' outweigh the 'costs of bias' as also the results from Rapach and Wohar (2004) suggest.[7]

While we follow the Behavioral Equilibrium Exchange Rate (BEER) approach by Clark and MacDonald (1998) to estimate 'fundamentally justified' values of specific currencies, a number of other concepts with similar acronyms have been proposed – most notably, the Fundamental Equilibrium Exchange Rate (FEER) approach by Williamson (1994a) and the Natural Real Exchange Rate (NATREX) approach by Stein (1995). Whereas the latter two concepts are normative and involve ad-hoc judgments on the size of central parameters, the BEER approach (which will be introduced in the next section) is rather statistical and free of normative elements. Due to the high number of publications dealing with equilibrium exchange rates and the vast number of competing concepts, a literature survey on equilibrium exchange rate would necessarily have to be very selective. Instead we refer to the excellent survey articles by MacDonald (2000), Driver and Wren-Lewis (1998), Driver and Westaway (2005) and Williamson (1994b, and 2009).

This paper is structured as follows: In section 2 we introduce the BEER concept and discuss the choice of fundamentals. In section 3 we present the data, its sources and how the data has been transformed. In section 4 we give a short overview of first and second generation panel unit root tests, which will then be applied – together with simple univariate unit root tests

[5]Methodologically similar, Baltagi et al. (1997 and 2000) compare the out-of-sample predictive performance of homogeneous versus various heterogeneous estimators for the demand for gasoline respectively cigarettes.

[6]The DOLS estimator accounts for potentially endogenous regressors as opposed to the simple OLS estimator. Not accounted for endogeneities would be another potential source of bias (simultaneity bias) we avoid by using DOLS.

[7]As a 'safety net' we additionally apply the *group-mean* DOLS estimator by Pedroni (2004), which gives consistent estimates of the *average* cointegration slopes, regardless whether the long-run parameters are homogeneous or heterogeneous across countries.

for each of the countries individually – to examine the order of integration of the involved time series. After a short methodological introduction, we will conduct the country-by-country estimations in section 5. We then turn to the panel analysis. In section 6 we introduce two kinds of panel cointegration tests, which will then be applied to test for long-run relationships among various subsets of the variables. In section 7 we conduct conditional forecasts over the hold-back-period to assess the relative forecast accuracy of both models, which may be regarded as an informal out-of-sample stability analysis. Based on the results of this exercise, we choose a preferred specification, which is then reestimated over the full sample to derive equilibrium real effective exchange rates and currency misalignments. Finally, section 9 concludes the article.

2 The BEER Concept and the Choice of Fundamentals

In this section we present the origins of the BEER concept and shortly present its theoretical foundations. For extensive surveys on the BEER concept and related concepts (such as CHEER, FEER, ITMEER, NATREX) we refer to Driver and Westaway (2005) and MacDonald (2000). The BEER concept arised from the discomfort with Purchasing Power Parity (PPP) as a reasonable explanation for the observed real exchange rate behavior. Shocks to the real exchange rate have been found to be too persistent to be accordant with PPP (the so called 'PPP-puzzle' (Rogoff 1996)), 'typical' half-life estimates to shocks to PPP being around four years. The very slow mean reversion speeds therefore provide at best support for an 'ultra-long-run' concept of equilibrium exchange rates (Edwards and Savastano, 1999). Sometimes even no significant mean reversion has been found.[8] From a theoretical point of view PPP has been criticized as a concept for the determination of the equilibrium exchange rate, because it ignores the role of capital flows and any real determinants of the real exchange rate (MacDonald, 2000). The aim of the BEER approach is on the one hand to be better able to relate the observed exchange rate behavior to movements in certain other variables, and on the other hand to correct the theoretical shortcomings of PPP as a method for determining equilibrium exchange rates. A general problem with respect to equilibrium exchange rates is that they are not observable. Consequently, estimated long-run relationships between the *observed* real exchange rate and the fundamentals are assumed to equal the equilibrium exchange

[8]Surveys can be found in MacDonald (2000), and Taylor and Taylor (2004).

rate towards which the real exchange rate adjusts (Égert, 2004).[9]

The BEER approach is derived from the real interest parity condition (thereby accounting for capital flows) and – assuming that the expected future real exchange rate is a function of fundamental variables – also takes into account real determinants.

Following, we briefly present the theoretical concept based on MacDonald (2000):

Starting point is the risk-adjusted real UIP condition (which is completely disregarded in the concept of PPP):

$$\Delta q_{t+k}^{e} = -(r_t - r_t^*) + \lambda_t \tag{1}$$

where q_t denotes the real exchange rate, e the expectations operator, r_t the ex-ante real interest rate and λ_t a risk premium. The real exchange rate is expressed in foreign currency units per unit of domestic currency, so that an increase represents a (real) appreciation of the domestic currency.

Rearranged for the real exchange rate we have:

$$q_t = q_{t+k}^{e} + (r_t - r_t^*) - \lambda_t \tag{2}$$

As the expected real exchange rate is unobservable, this relationship is hard to test empirically; q_{t+k}^{e} is therefore interpreted as the fundamental or long-run component of the real exchange rate, denoted $\overline{q}_t$. Substituting into (2) yields:

$$q_t = \overline{q}_t + (r_t - r_t^*) - \lambda_t \tag{3}$$

In a popular stock-flow consistent model of exchange rate determination, Faruqee (1995) identifies the stock of net foreign assets and a set of exogenous variables as fundamental determinants of the equilibrium (or long-run) exchange rate. In an extension of this model, Alberola et al. (1999) decompose the real exchange rate into an external and an internal real exchange rate, both of which relate to one specific theory of exchange rate determination. The external rate is defined as the ratio of the price of domestic relative to foreign tradable goods and is connected to the notion of external balance, whereas the internal rate is defined as the relative price of non-tradable goods to tradable goods within each country and connected to the notion of internal balance. The equilibrium exchange rate is the one at which external and internal balance are achieved, i.e. the tradable goods market is cleared

[9] Funke (2005) describes the BEER as the 'data-determined systematic component of the exchange rate in the medium and long run'.

(and the desired net foreign asset position achieved), and there is no excess demand for non-tradable goods.[10]

We will not go into any model-specific details here, but instead give some intuitive motivation for the inclusion of specific fundamentals, which are hypothesized to account for the time-varying value of the real exchange rate, and which are theoretically founded by one model or another or are commonly used in the literature in an ad-hoc style. For model-specific derivations we refer to Frenkel and Mussa (1985), Faruqee (1995), Chinn and Johnston (1996), and Alberola et al. (1999).

Net Foreign Assets (+)

The effect of a country's net foreign asset position on the equilibrium exchange rate is expected to be positive. According to standard intertemporal macroeconomic models, a higher stock of net foreign *liabilities* causes interest payments to increase, which ultimately have to be paid for by improved trade balances. In order to generate higher net exports, the competitiveness of the respective country has to improve, necessitating a depreciation of the real exchange rate.

Productivity Bias (+)

According to the well-known Balassa-Samuelson (BS) hypothesis, relatively larger increases in productivity in the tradable goods sector compared to the non-tradable goods sector are connected with a real appreciation of the domestic currency. The rise in relative productivity of the tradable goods sector causes wages in the tradable goods sector to increase. Wage equalization across the sectors ensures that wages in the non-tradable sector also increase (which is not compensated for by an accordant rise in productivity). Consequently, the overall price level is higher and we observe a real appreciation of the equilibrium exchange rate.

Government Consumption (+/-)

According to Genberg (1978), Bergstrand (1991), and MacDonald (1998) the presence of non-traded goods can furthermore introduce a demand side bias, if the income elasticity of demand for non-traded goods is

[10] The BEER approach therefore includes the same set of fundamentals as the FEER approach. However, the latter implies the calibration of a sustainable current account and therefore follows a normative approach.

greater than 1. As government consumption is primarily devoted to non-traded goods (compared to private consumption), a rise in government consumption (or a redistribution of income towards the government) therefore increases the relative price of non-tradable goods and causes the equilibrium exchange rate to appreciate.[11] On the other hand, a growing budget deficit might also cast doubt on the sustainability of fiscal policy, destabilize the economy and lead to a real depreciation (Melecký and Komárek, 2007).

Terms of Trade (+/-)

When traded goods are imperfect substitutes, their relative price might change due to changes in supply and demand caused by changes in underlying determinants such as consumer preferences or differing growth rates (Edwards, 1989 and Nilsson, 2004). Import and export price indices will be affected differently across countries, thereby changing the relative terms of trade. The effect of terms of trade shocks to the real exchange rate is however ambiguous (Melecký and Komárek, 2007). On the one hand, an increase in the (relative) prices of a country's export goods gives rise to a positive substitution effect, because domestic producers shift their production towards tradable goods. This will cause wages in this sector to increase relative to the ones in the non-tradable goods sector. If wages subsequently equalize across the sectors, this will drive up the overall price level and thereby lead to a real appreciation. However, at the same time a positive wealth effect (reflected in the improved current account) may generate higher demand for non-tradable goods and necessitate a real depreciation to restore internal balance (Melecký and Komárek, 2007).

Openness (-)

Openness is often introduced as a proxy for trade liberalization. A higher degree of openness (i.e. a removal of trade restrictions) may lower domestic prices and cause a real depreciation (Goldfajn and Valdes, 1999 and Elbadawi, 1994). As emphasized by Dufrenot and Yehoue (2005) this variable is however more likely to be relevant for developing or emerging countries than for industrial ones, since we would not expect the degree of openness to vary a lot for these countries. We

[11] See Chinn (1997) for a formal implementation of this argument in a real exchange rate model.

therefore think that its influence should, with the exception of China, be fairly limited in our set of countries.

3 Data Description

Our panel consists of the US and 16 of its major trading partners: Australia, Belgium, Canada, China, France, Germany, Ireland, Italy, Japan, Korea, Mexico, Netherlands, Spain, Sweden, Switzerland and the UK. These countries either make up a a significant part of US trade, or are an issuer of a major currency, i.e. their currency circulates widely outside the country of issue. Some countries have not been included in the panel due to too much missing data for some of the relevant variables (Taiwan, Singapore, Hong Kong, Malaysia, and Brazil). The sample consists of quarterly data from 1986Q1 to 2006Q4. For some variables and countries, data interpolation techniques are applied to obtain quarterly data. All variables except for the net foreign asset to GDP ratio, the trade balance to GDP ratio and the openness ratio are measured relative to the trade-weighted average of the respective variables of the trading partner countries. Trading weights are constant throughout the sample and are taken from Bayoumi et al. (2005).[12] The trade-weights for all included countries (rescaled so that they sum to 1) are presented in table 9. Although the countries have been chosen with special regard to the US, table 1 shows that all countries included in the panel make up a large proportion of trade for each of the other countries (ranging from 69 percent for Sweden to 90 percent for Mexico). So we believe that the results are meaningful for all countries.

Table 1: Sum of Trade Weights of Partner Countries

AUS	BEL	CAN	CHN	FRA	GER	IRL	ITA	JPN
78%	89%	75%	78%	70%	84%	74%	70%	72%
KOR	MEX	NLD	ESP	SWE	CHE	UK	US	
74%	90%	77%	77%	69%	80%	77%	76%	

Data is taken from IMF International Financial Statistics (IFS) unless stated differently. Exchange rates are defined in indirect quotation so that an increase in the exchange rate implies an appreciation of the domestic currency. Annual Consumer Price Inflation (CPI) for China has been obtained from the

[12] Reference period for the calculation of these weights is 2002/2003 and weights take into account direct bilateral competition as well as third-market competition.

National Bureau of Statistics of China, interpolated to quarterly frequency and updated with growth rates from IFS.[13] The real effective exchange rate ***(reer)*** is calculated as the geometric average of the 16 bilateral nominal exchange rate indices multiplied with the respective CPI ratio. Data on net foreign assets is taken from Lane and Milesi-Ferretti (2007) and updated by accumulating current account balances. To scale the variables, we divide by GDP to obtain the net foreign asset to GDP ratio ***(nfa)***. Alternatively, we also use the trade balance to GDP ratio ***(tb)***, which is available at a higher frequency.[14] For China and Ireland annual GDP data is converted to quarterly frequency for some periods.[15] Relative terms of trade ***(tot)*** are defined as the ratio of the domestic export unit value and the domestic import unit value divided by the geometrically weighted (using the same weights as above) foreign ratios. The relative productivity differential ***(ntt)*** is proxied by the ratio of the domestic consumer price index (CPI) to the domestic wholesale price index (WPI) divided by the geometrically weighted foreign ratios. The real interest rate is calculated as the difference between the nominal long-term interest rate and the percentage change of the CPI index compared to its value in the same quarter in the year before. Nominal long-term interest rates are partly taken from sourceOECD. For China the lending rate is used instead, and for Mexico the 3-month treasury bill rate is taken until 1989Q4.[16] The real interest rate differential ***(rird)*** is then calculated as the difference between the domestic real interest rate and the trade-weighted arithmetic average real interest rate of the trading partners. Relative government consumption ***(gov)*** is defined as the domestic government consumption to GDP-ratio relative to the trade-weighted ratios of government consumption to GDP-ratios of the trading partners. Openness ***(open)*** is defined as the ratio of the sum of exports and imports (in absolute values) to GDP. While *reer*, *gov* and *tot* are expressed in logarithmic terms, *nfa*, *tb*, *open* and *rird* are measured as fractions of 100.

[13] We use the cubic-spline interpolation technique. However, results are robust to changing the interpolation technique.

[14] *tb* has first been introduced as an alternative for *nfa* by MacDonald and Dias (2007). A positive long-run relationship between *nfa* and *reer* implies a negative long-run relationship between *tb* and *reer*.

[15] For Ireland until 1996Q4, for China until 1998Q4.

[16] One missing value for Mexico for 1986Q3 is proxied by the average of the preceding and the succeeding quarter.

4 (Panel) Unit Root Tests

As noted in the introduction we put ourselves once fully in the position of a time series econometrician, and once in the position of a panel econometrician. We also follow this approach when assessing the order of integration of the involved time series.

Time series econometricians usually apply the familiar Dickey-Fuller type univariate unit root tests to assess the order of integration of certain time series. Because the tes ts are fairly standard we do not introduce them here. We conduct two types of tests, the well-known augmented Dickey Fuller (ADF) test as well as the Dickey-Fuller generalized least squares (DFGLS) test proposed by Elliott et al. (1996).[17] To save space, test results are not reported here, but can be obtained from the author upon request. Later, we will conduct country-by-country tests for cointegration for all different subsets of variables, in which only variables are included that are tested to be I(1) in the country-by-country unit root tests.

Univariate unit root tests are known to have low power to reject the null hypothesis of a unit root, if the respective time series contains a 'near unit root', i.e. the autoregressive parameter is close to but lower than one. Increasing the dimension of the panel from $N = 1$ (where N is the number of cross section units) to $N > 1$ lowers the probability of committing a type 2 error. However, a number of complications arise in the presence of more than one cross section unit. This is reflected in the large number of panel unit root tests which have been developed. Whereas the rejection of the null hypothesis of a unit root is trivial to interpret if $N = 1$, it is less obvious for $N > 1$. Rejection could either mean that the respective series is stationary for all cross units of the panel, or that it is stationary for a fraction of the cross units included in the panel only.

We implement three panel unit root tests: The Levin-Lin-Chu (2002) (LLC) test, the Im-Pesaran-Shin (2003) (IPS) test, and the Pesaran (2007) CIPS (cross-sectionally augmented IPS) test. We will only comment on the main differences among these three panel unit root tests. For a comprehensive formal treatment of these and other panel unit root tests (including the derivation of the tests and test statistics, and some Monte Carlo evidence) see Breitung and Pesaran (2005), which we also draw upon in this short survey. For an overview of second generation panel unit root tests see Hurlin and Mignon (2007).

The tests we consider differ in two main aspects: first, whether a common

[17] Elliot et al. (1996) and Ng and Perron (2001) show that this test is more powerful when an unknown trend and/or mean is present.

autoregressive process is assumed under the alternative hypothesis, and secondly, how cross-sectional dependencies are dealt with. The LLC and the IPS test belong to the so-called 'first generation' of panel unit root tests, while the Pesaran CIPS test is one of several recently proposed 'second generation' tests. The difference between first- and second generation tests is that the latter take into account cross-sectional dependencies, while the former do not (or only to a very limited extent by including common time dummies in the test regressions). While the LLC and IPS tests are still the most commonly applied, several studies have shown that the empirical size of a panel unit root test can vastly diverge from the nominal size if cross-sectional dependencies are disregarded (see for instance Banerjee et al., 2005).

Before we more carefully address the issue of cross-sectional dependencies, we have a closer look at the LLC and IPS tests. The LLC test starts from the following *pooled* ADF test regression (lags of the differenced dependent variable are disregarded for simplicity):

$$\Delta y_{i,t} = \alpha_i + \delta_i t + \theta_t + \rho_i y_{i,t-1} + \varepsilon_{i,t}, \tag{4}$$

The test allows for some degree of heterogeneity by including country fixed effects (α_i), country-specific deterministic (time) trends ($\delta_i t$) and *dampens* the effects of cross-sectional dependencies by including common time dummies (θ_t). However, ρ_i is obtained by running a *pooled* OLS regression, i.e. $\rho_i = \rho$ for all countries. The null and alternative hypotheses of the LLC test are $H_0 : \rho_i = 0 \forall i$, respectively $H_a : \rho_i = \rho < 0 \forall i$. We see that the LLC test is very restrictive under the alternative hypothesis by assuming that the series follows the same autoregressive process for all cross units. This 'drawback' is addressed by the IPS test, which is a *group-mean* test. The alternative hypothesis is that $H_a : \rho_i < 0$ for at least one i. In contrast to the LLC test, where the estimate of ρ is obtained by running a pooled regression, ρ_i is estimated for each cross unit individually before a group-mean test statistic, $\bar{t}$, is obtained though averaging the individual t-statistics. The test is less restrictive because ρ_i may be different for different i under H_A. Both tests considered so far can only handle cross-sectional dependencies to a very limited degree (by including common time dummies). However, this approach is not suitable if the pair-wise cross-section covariances of the error terms are different across the individual series (Pesaran, 2007). Pesaran (2007) therefore proposes a generalized IPS test, which allows a common factor to have different effects on each cross unit.[18]

[18]Banerjee et al. (2005) show that the empirical size of panel unit root tests is much higher than the nominal level in case there are cross-unit cointegration relationships which are not taken into account in the respective critical values. In other words, it is likely that

It is based on the following cross-sectionally augmented ADF (CADF) test regression(s)

$$\Delta y_{i,t} = \alpha_i + \delta_i t + \rho_i y_{i,t-1} + \gamma_i \bar{y}_{t-1} + \varsigma_i \Delta \bar{y}_t + \varepsilon_{i,t}, \tag{5}$$

where Δy_t and the lagged cross-sectional averages of y_t serve as proxies for the effects of an unobserved common factor (Pesaran, 2006 and 2007). After having run the CADF regressions for each of the cross units individually, a group-mean statistic is again calculated (analogously to the IPS test), which can then be compared to the respective critical value.

According to the results from the (pooled) LLC test and the (group-mean) IPS test presented in table 2, all but two variables are integrated of order one.[19] Only *tot* and *rird* are tested to be stationary processes (at least for a non-zero fraction of the countries included in the panel) at the 5% level.

To determine the (approximately) appropriate lag order for the CADF regressions underlying the CIPS test we run auxiliary ADF test regressions for each of the cross-section units prior to the estimation of the CADF test regressions. We choose the lag order that maximizes the Hannan-Quinn criterion (HQC) allowing for a maximum lag length of 8. Afterwards we apply the CIPS test based on CADF-regressions with the respective previously determined lag-lengths.

Table 3 shows the CIPS test results together with the selected average lag lengths. The results show that our prior classification regarding the order of integration of the involved series is still valid if we (additionally) account for cross-sectional dependencies. This result together with the uniform results obtained from the LLC and IPS test suggest that the classification is robust.

Summarizing, all series apart from *tot* and *rird* are I(1) based on panel unit root tests. As *reer* has also been tested to be I(1), the potential explanatory variables need to be I(1) as well in order to avoid unbalanced equations. This implies that we will test for long-run relationships among subsets of

---

the true null of non-stationarity is rejected too often (i.e. the probability of a type 1 error is higher) in the presence of cross-unit cointegration relationships. As an example, they show that evidence based on panel unit root tests in favor of relative PPP collapses once cross-section cointegration is taken into account by using suitable critical values.

[19]For the series in levels we include individual intercepts and individual trends, whereas we include only country-specific intercepts for the series in first differences. We first started with all 17 countries included in the panel. Results for the order of integration of *reer* were borderline however. Country-specific individual unit root test results revealed that the *reer* of Mexico and Switzerland seem to be (trend)stationary. We excluded these countries from the subsequent analysis because of the supposedly different order of integration of *reer*. Applying the panel unit root tests to the smaller panel of 15 countries we obtain the above-described clear results.

Table 2: First Generation Panel Unit Root Test Results

		REER	NFA	NTT	TB
LLC	level	-0.67	-0.08	-0.33	1.87
	1st diff.	-21.30^{***}	-2.48^{***}	-19.88^{***}	-30.87^{***}
IPS	level	-1.20	0.05	1.57	-1.21
	1st diff.	-20.76^{***}	-12.62^{***}	-19.43^{***}	-32.84^{***}
		TOT	GOV	OPEN	RIRD
LLC	level	-1.86^{**}	0.84	2.79	-2.18^{**}
	1st diff.	-30.33^{***}	-33.28^{***}	-21.22^{***}	-21.94^{***}
IPS	level	-4.10^{***}	1.50	1.65	-3.20^{***}
	1st diff.	-29.52^{***}	-32.76^{***}	-23.06^{***}	-26.69^{***}

Note: Reported are the $\bar{t}$-statistics for the IPS test and the t^*-statistics for the LLC test. We allow for individual deterministic trends and constants, and series are demeaned. ***, **, and * denote significance at the 1, 5, and 10% level, respectively.

Table 3: CIPS Second Generation Panel Unit Root Test Results

		REER	NFA	NTT	TB
CIPS	level	-2.58	-2.37	-2.29	-2.36
	avg. lags	1.53	3.13	1.00	1.00
CIPS	1st diff.	-5.92^{***}	-3.23^{***}	-5.80^{***}	-5.877^{***}
	avg. lags	0.47	2.13	0.33	0.60
		TOT	GOV	OPEN	RIRD
CIPS	level	-2.99^{***}	-2.03	-1.70	-2.78^{**}
	avg. lags	0.80	0.47	0.73	1.87
CIPS	1st diff.	-6.11^{***}	-5.94^{***}	-5.66	-5.62^{***}
	avg. lags	0.60	0.27	0.73	1.87

Note: We report the $\bar{t}$-statistics. Avg. lags denotes the average lag length of the underlying CADF test regressions. ***, **, and * denote significance at the 1, 5, and 10% level, respectively.

reer, nfa, tb, ntt, gov and *open* only, and disregard *tot* and *rird* in the panel cointegration analysis. We however do not generally disregard these variables in the country-by-country cointegration analyzes. Based on the country-by-country univariate unit root tests *tot* appears to be non-stationary for some of the countries. This result contrasts with the IPS and CIPS-test results, which maintain that *tot* is not stationary for any of the cross units. It may be regarded as a typical example for the previously noted lower power of individual unit root tests compared to panel unit root tests in rejecting a near unit root. However, as we consider the situation of a time series analyst who only takes into account information based on regressions for a single cross unit, this knowledge would not be available to him. In line with the previous argument, the fraction of nonstationary series is larger according to the country-by-country unit root tests. Consequently, the number of specifications which need to be tested for cointegration is also larger compared to the panel analysis.

5 Country-by-Country Analysis

5.1 Tests for Cointegration and Estimation Methodology

We (primarily) apply the DOLS estimator to estimate the long-run relationships among subsets of variables. We choose this specific estimator for two reasons: First, it can easily be applied in (nonstationary) panel regressions as well. By choosing the equivalent panel estimator we narrow down the reasons for possibly different results, because different estimation results cannot be attributed to a different estimation technique then. Secondly, the DOLS estimator accounts for potential endogeneities among the variables.[20]

DOLS has been introduced by Saikkonen (1991) and Stock and Watson (1993) and extended to panel analysis by Kao and Chiang (1997). By incorporating leads and lags of the first differences of the regressors endogenous feedback effects from the dependent variable to the regressors are absorbed. In contrast to OLS, the DOLS estimator is therefore consistent, even if regressors are endogenous. Without a cross-sectional dimension, a DOLS(k_1,k_2) regression can be written as:

$$Y_t = \beta_0 + \sum_{i=1}^{n} \beta_i X_{i,t} + \sum_{i=1}^{n} \sum_{j=-k_1}^{k_2} \gamma_{i,j} \Delta X_{i,t-j} + \varepsilon_t,$$

[20]The regular OLS estimator is biased if regressors are not weakly exogenous.

where k_1 and k_2 denote the numbers of leads and lags, which can be chosen according to various information criteria.

To check which of the estimated regressions are spurious and which form long-run relationships, we perform Engle-Granger cointegration tests. Critical values for the Engle-Granger cointegration tests are calculated from the response surface function presented in MacKinnon (1990). Using these obtained critical values we check whether the residuals from the supposed long-run relationship, $\upsilon_t = Y_t - \beta_0 - \sum_{i=1}^{n} \beta_i X_{i,t}$, are indeed stationary. If not, then the null of no cointegration cannot be rejected and the relationship is spurious.

In order to check whether the results are robust against using another estimation technique, we additionally present results obtained from fully modified OLS (FMOLS) regressions. The FMOLS estimator also accounts for endogenous regressors, but in contrast to the DOLS estimator, the bias correction is obtained in a non-parametric way.[21]

5.2 Estimation Results

So far we have neither presented country-by-country unit root test nor cointegration test results, since this would consume too much space due to the large number of series/specifications. Instead, we directly report the DOLS and FMOLS estimation results in table 2.11 for the specifications for which the following three conditions are fulfilled: First, all series included in each of the specifications have to be I(1) according to the *individual* countries' unit root test results, secondly, the residual of the cointegration vector has to be stationary, and thirdly, there has to be mean reversion in the error correction model (ECM) representation so that the regression can be interpreted as a real exchange rate equation. Consequently, for the large number of remaining specifications which are not reported here, at least one of the three conditions is not fulfilled. Only DOLS(1,1) results are reported, because the inclusion of further leads and lags changed the point estimates only marginally. The reported standard errors of the DOLS regressions are heteroscedasticity and autocorrelation consistent (HAC).

Overall, the results are very mixed. First, for four countries (AUS, ESP, FRA, and UK) we do not find any specification fulfilling all the above criteria. Secondly, we do not find a single specification (containing the same set of explanatory variables) for the remaining countries, either pointing towards true heterogeneity of the different BEER equations or to imprecise

[21] For an intuitive derivation of the FMOLS estimator see Patterson (2000).

estimates due to the relatively low number of observations for each of the cross-section units. Stationary combinations including *nfa* are only obtained for four countries (CHN, JPN, IRE, and NLD) and the estimated signs are counterintuitive.[22] Also *tb* is only included and significant in a very few cases (CAN, CHN, and US).[23] Based on the country-specific estimation results we could therefore doubt their economic significance. However, we abstain from drawing further conclusions at this point, and will instead reconsider the role of *nfa* and *tb* in the context of the subsequent panel analysis.

Whereas evidence in favor of these commonly hypothesized determinants is at best weak, we find strong evidence in favor of the BS effect. It is included in all specifications and the sign and size of the coefficient estimates are plausible (a one percent increase in *ntt* is connected with a 0.65 percent to 2.67 percent appreciation of the domestic currency).

Evidence for the other determinants is mixed again. *tot* is included in only three specifications (CAN, CHN, and SWE), which is not surprising against the background of the panel unit root tests, according to which *tot* is stationary, whereas *reer* is nonstationary. For the above three specifications the estimated coefficient of *tot* is always positive thereby providing weak evidence in favor of the substitution effect dominating the wealth effect. While *open* is included in the specifications of six countries (CAN, GER, JPN, KOR, NLD, and SWE) the respective point estimates vastly differ (estimated coefficients range from -0.35 for CHN to about -6.43 for JPN).

gov has a significantly negative impact on *reer* in three countries (CAN, CHN, and IRL). This lends some support to the hypothesis that concerns about fiscal sustainability dominate the effects of (relatively) higher government spending on non-traded goods in these countries. In the case of JPN, the relationship among *gov* and *reer* is significantly positive.

The FMOLS results are in most cases very similar to the DOLS(1,1) results. Consequently, our results do not seem to depend on the choice of the estimator.

In the last column of table 2.11 (labeled α), we report the estimated adjustment coefficient in each respective ECM. Based on these values we calculate the implied half-life time of deviations from the estimated equilibrium exchange rate for each of the specifications according to $t^{1/2} = \frac{ln(0.5)}{ln(1+\alpha)}$, where α is the respective adjustment parameter in the ECM.

The half-life time varies between less than 2 quarters and about 5 quarters for most of the specifications, which is a dramatic improvement over

[22] For JPN there is a smaller specification excluding *nfa* which is also stationary. So *nfa* does not seem to be necessary to 'achieve' stationarity.

[23] The significance depends on whether results are obtained by DOLS or FMOLS.

previously estimated half-life times implied by PPP, and an even larger 'improvement' compared to our own finding that *reer* is I(1) implying that there is no significant mean reversion (and therefore no evidence supporting PPP) at all.[24] Notwithstanding the favorable results in terms of adjustment speeds towards equilibrium, the obtained country-by-country results are very mixed in general, and one should therefore not overemphasize these findings, because they may be based on imprecise estimates.

Summarizing the country-by-country results, it is highly difficult to find sensible and robust specifications for the regarded countries. This may either be attributed to the low power of cointegration tests in small samples (68 observations), and, closely related, to imprecise estimates, or, trivially, to the non-existence of sensible cointegration relationships among the variables. However, there is one exception: We obtain sensible results with respect to *ntt*, which is estimated quite robustly across units with the expected positive sign and by and large reasonable coefficient estimates. This lends strong support to the BS-effect.

6 Panel Analysis

6.1 Tests for Panel Cointegration and Estimation Methodology

There is a close analogy between panel cointegration tests and panel unit root tests. Some of the tests are based on group-mean estimates, others on pooled estimates. Some take into account cross-sectional dependencies, while others do not. We will apply two representative (bundles of) panel cointegration tests: the very popular Pedroni (2004) test(s) for panel cointegration and the recently introduced test(s) by Westerlund (2007). A comprehensive survey on panel cointegration tests is provided by Breitung (2005).

Since the Pedroni panel cointegration test (2004) is residual-based, it can be regarded as a panel equivalent of the Engle-Granger test for cointegration commonly applied in time series analysis. Pedroni proposes seven tests, of which three are group-mean tests and the remaining four are pooled tests (with the respective differing alternative hypotheses). A detailed discussion of each individual test statistic is outside the scope of this paper and we refer to Pedroni's (2004) original article for further details. Similarly as in the case of the Johansen test for cointegration, short-run parameters and country-specific deterministic trends are filtered out in two first stage regressions.

[24]Only in GER and ITA adjustment to shocks is much slower (the implied half-life time is about 3 years for Germany and almost 6 years for Italy).

By doing so the Pedroni test allows for country-specific short-run effects and different lag-lengths in the test-regressions (in contrast to the formerly heavily applied test by Kao, 1999). In general, it can be regarded as a sign of robustness if several of the different test statistics lead to the same test decision, because evidence based on Monte Carlo simulations has shown that the various test statistics perform differently depending on the panel dimension and the assumed data generating process.

The error-correction based test by Westerlund (2007) does not only allow for various forms of heterogeneity, but also provides p-values which are robust against cross-sectional dependencies via bootstrapping.[25] In short, it is tested whether the null of no error correction can be rejected (either for the whole panel or for a non-zero fraction of the cross units depending on whether a pooled or group-mean estimation is performed). If the null can be rejected, there is evidence in favor of cointegration.

While two of the four tests are panel tests with the alternative hypothesis that the whole panel is cointegrated ($H_A^P : \alpha_i = \alpha < 0$ for all i), the other two tests are group-mean tests which test against the alternative hypothesis that for at least one cross-section unit there is evidence of cointegration ($H_A^G : \alpha_i < 0$ for at least one i). For the group-mean test statistics, the error correction coefficient is estimated for each cross-section unit individually, and then two average statistics (denoted G_t, respectively G_α) are calculated.[26] In the pooled tests, the series of each cross-section unit are 'cleaned' first (of dynamic nuisance parameters, unit-specific intercepts and/or trends), before the conditional (or 'cleaned') panel error correction model is estimated to obtain a common α estimate, which is checked for significance.

6.2 Estimation Results

While we apply the Pedroni test to search for long-run relationships among twelve different subsets of variables, the Westerlund test is only applied to the specifications for which the Pedroni tests provide strong evidence in favor of cointegration.[27] The number of subsets is determined by the results of the panel unit root tests, according to which only *reer, nfa, tb, ntt*, and *open* are

[25]For a description of the respective STATA procedure see Persyn and Westerlund (2008).

[26]For more details on the test-statistics and their derivation see the above reference.

[27]Banerjee et al. (2004) show that panel cointegration tests can be largely oversized in the presence of cross-unit long-run relationships. Not accounting for such relationships therefore makes it more likely to obtain a finding in favor of cointegration, which may be false. It is therefore sensible to apply the Westerlund test which accounts for cross-sectional dependencies only to the specifications for which the Pedroni test points towards cointegration.

Table 4: Pedroni Panel Cointegration Test Results

Specification	*M1*	*M2*	*M3*	*M4*	*M5*	*M6*
	REER	REER	REER	REER	REER	REER
	NFA	NFA	NFA	NFA		
					TB	TB
	NTT	NTT	NTT	NTT	NTT	NTT
		GOV		GOV		GOV
			OPEN	OPEN		
Panel Tests						
ν-stat.	1.13	1.16	0.92	1.04	1.14	0.68
ρ-stat.	0.00	0.37	0.56	0.78	−0.79	0.03
t-stat. (ADF)	−0.80	−0.59	−0.47	−0.42	−1.66	−1.05
t-stat. (PP)	-1.35^{*}	−1.12	−0.83	−0.02	−2.32	−0.92
Group Mean Tests						
ρ-statistic	1.26	1.43	1.63	1.75	0.59	1.45
t-stat. (PP)	0.04	0.04	0.21	0.16	−0.79	−0.06
t-stat. (ADF)	−1.11	−0.96	−0.75	0.78	-2.34^{***}	−0.75

Specification	*M7*	*M8*	*M9*	*M10*	*M11*	*M12*
	REER	REER	REER	REER	REER	REER
	TB	TB				
	NTT	NTT	NTT	NTT	NTT	NTT
		GOV		GOV		GOV
	OPEN	OPEN			OPEN	OPEN
Panel Tests						
ν-stat.	1.04	0.90	1.67^{**}	1.05	1.12	0.86
ρ-stat.	0.15	0.65	-1.68^{**}	−0.63	−0.47	0.01
t-stat. (ADF)	−0.80	−0.56	-2.18^{**}	-1.45^{*}	−1.12	−0.95
t-stat. (PP)	-1.35^{*}	−0.97	-2.40^{***}	-1.65^{**}	-1.31^{*}	−0.87
Group Mean Tests						
ρ-statistic	1.24	1.72	−0.08	0.61	0.73	1.14
t-stat. (PP)	−0.14	0.08	−1.26	−0.69	−0.36	−0.19
t-stat. (ADF)	−1.12	−1.17	-1.86^{**}	-1.57^{*}	−0.76	−0.62

Note: ***,** and * denote the significance levels of 1, 5, and 10%.

Table 5: Westerlund Panel Cointegration Test

Statistic	Value	Z-value	p-value	Robust p-value
Gt	-2.515	-3.180	0.001	0.010
Ga	-9.452	-1.644	0.050	0.030
Pt	-7.708	-2.114	0.017	0.098
Pa	-7.580	-2.922	0.002	0.023

Note: Note: Optimal lag/lead length determined by Akaike Information Criterion with a maximum lag/lead length of 3. Width of Bartlett-kernel window set to 3. We allow for a constant, but no deterministic trend in the cointegration relationship. Number of bootstraps to obtain bootstrapped p-values which are robust against cross-sectional dependencies set to 400.

I(1), and additionally by our own requirement that *ntt* is included in all tested subsets. While this may seem somewhat arbitrary, we believe it is sensible to do so against the background of the country-by-country estimation results providing strong evidence of the BS-effect.

One result is particularly remarkable: Evidence in favor of cointegration is the strongest when only *ntt* is considered as a regressor (see table 4). In this case, 5 of the reported 7 statistics point towards the presence of a cointegrating relationship among the variables. If the specification additionally includes *gov* we can only reject the null of no cointegration in 3 of 7 cases. For three other specifications only one of the statistics points towards cointegration, for the remaining 7 specifications we find no evidence of cointegration at all, although the former specifications are restricted versions of the latter. In table 5 we present the Westerlund test results for *M9*. In the last column we present the bootstrapped p-values, which account for cross-sectional dependencies. They point towards *reer* and *ntt* being cointegrated, thereby supporting the overall Pedroni test result. According to three of the four test statistics we can reject the null of no significant error correction at the 5%-level, while one rejects only at the 10% level.[28]

As stated above, we also focus on the DOLS estimator in our panel analysis. First we run *pooled* DOLS regressions allowing for two-way fixed effects (country fixed effects and common time effects). The pooled is estimator

[28] According to the asymptotic p-values in the second-last column, we can reject the null of no cointegration in all cases.

Table 6: Pooled DOLS Estimates

	M1	*M5*	*M7*	*M9*	*M10*	*M11*
NFA	-0.09^{***}					
	(0.02)					
TB		-0.83^{***}	-0.29^{**}			
		(0.08)	(0.11)			
NTT	1.15^{***}	0.98^{***}	0.75^{***}	1.18^{***}	1.31^{***}	0.76^{***}
	(0.10)	(0.10)	(0.08)	(0.10)	(0.09)	(0.08)
GOV					-0.32^{***}	
					(0.07)	
OPEN			-0.38^{***}			-0.43^{***}
			(0.06)			(0.05)
Poolability				10.56^{***}	15.51^{***}	

Note: Driscoll and Kraay standard errors in brackets. ***,** and * denote the significance levels of 1, 5, and 10%, respectively. The null hypothesis of the Roy-Zellner test is poolability/slope homogeneity across countries.

is only unbiased if the cointegration slopes are equal across countries. We therefore formally test whether the poolability hypothesis is not rejected after having performed the estimations.

We report Driscoll and Kraay (1998) standard errors, which account for within-group correlation, heteroscedasticity and cross-sectional correlation.[29] Because the inclusion of common time effects does not substantially change the point estimates of the other coefficients, we only present the point estimates without common time dummies included.

DOLS(1,1) results for the restricted sample are presented in table 6 for the specifications, for which at least one of the reported Pedroni test statistics is in favor of rejecting the null of no cointegration, however, given the results of the panel cointegration tests we focus on specifications 9 and 10 in the subsequent analysis. The findings suggest that only *ntt* is necessary to achieve stationarity.[30] The obtained point estimate of *ntt* is significantly positive and of reasonable size (with a point estimate of 1.18). Accordingly, the panel results support the results obtained in the country-by-country analysis with respect to the relevance of *ntt*.

[29]We used the STATA module 'xtscc' by Hoechle (2007) to provide robust standard errors.

[30]It contrast to all other combinations *reer* and *ntt* form a so-called irreducible cointegration relationship (Davidson, 1999).

Table 7: Group Mean DOLS and FMOLS Estimates

	M1		*M9*		*M10*	
	DOLS	FMOLS	DOLS	FMOLS	DOLS	FMOLS
NFA	-0.02	0.02				
	(-0.64)	(-0.20)				
NTT	1.43^{***}	1.42^{***}	1.24^{***}	1.21^{***}	1.31^{***}	1.42^{***}
	(13.07)	(12.57)	(16.38)	(14.16)	(18.73)	(14.70)
GOV					-0.30^{***}	-0.30^{***}
					(-2.93)	(-2.66)

Note: ***,** and * denote the significance levels of 1, 5, and 10%, respectively.

However, we have to be cautious in interpreting the results, because we have to reject the null hypothesis of poolability across countries according to the Roy-Zellner test (Roy, 1959, Zellner, 1962, and Baltagi, 2005), which is reported for specifications 9 and 10 at the bottom of table 6.[31]

Baltagi (2005) and Baltagi and Griffin (1997) suggest to choose a pragmatic approach. Instead of disregarding the pooled model if the poolability restriction is rejected, they propose to base the decision of whether pooling is advantageous or not on the out-of-sample forecast performance of the heterogeneous models vs. the pooled model. We will follow their advice.

Before we do so in the subsequent section, we provide group-mean DOLS and FM-OLS results of the cointegration slopes for three different models (*M1*, *M9*, and *M10*). Group-mean DOLS and FMOLS estimators have been introduced by Pedroni (2004). The advantage of these estimators is that they provide consistent estimates of the *average* cointegration slopes even if the slopes are in fact different across countries. We choose the above models for two reasons: First, evidence of cointegration has been strongest for *M9* and *M10*. Secondly, because *nfa* has commonly been included in BEER equations we want to check whether its numerically small coefficient estimate obtained in the pooled regression may be due to omitted variable bias in *M1*.[32] Estimation results are presented in table 7.

We observe that the estimate of the cointegration slope coefficient of *ntt*

[31] We only test the slope coefficients of the original regressors in levels (*ntt*, respectively *ntt*, *gov*) for homogeneity across countries, not the ones from the lags and leads of their first differences. See Vaona (2008) on how to implement this test in STATA.

[32] By pooling the data, we implicitly introduce this form of bias if the slope homogeneity restriction does not hold.

hardly changes in our preferred specification *M9* (from 1.18 in the pooled DOLS regression to 1.24 and 1.21 in the group-mean FMOLS and DOLS regressions, respectively). The same holds for *M10*, where the coefficient of *gov* either does not change much if we use the group mean DOLS or FMOLS estimator instead of the pooled estimator (-0.32 compared to -0.30). It is furthermore notable that the coefficient estimate of *nfa* is even smaller and insignificant in the group mean regressions (see *M1*). Consequently, the small coefficient estimate of *nfa* in *M1* in the pooled case may even be too large. Based on these findings and the country-by-country estimation results, *nfa* does not seem to be a fundamental determinant of *reer*.[33] Based on our results, it is therefore hard to argue that the depreciation of the USD in current years is a direct response to the growing US external liabilities. However, this surely does not mean that *nfa* does not have any influence on the *reer* at all. If one is on the one hand willing to accept the theoretically appealing proposition that net foreign liabilities ultimately have to be repaid and on the other hand empirical evidence in favor of a linear long-run relationship is at most weak, this may raise concerns of a (non-linear) sudden adjustment once certain thresholds are reached. We leave such a threshold-analysis as well as a robustness check of our results with respect to other countries and samples to further research.

7 Conditional Forecasts

Based on the results of the panel cointegration tests, the plausibility of the coefficient estimates, and the limited impact of other variables on *reer*, panel specifications *M9* and *M10* are the only panel specifications which we consider in our conditional forecasting exercise over the reserved part of the sample (2003Q1 to 2006Q4). Their forecasting performances are compared with the ones from the fully country-specific time series models. To assess the relative forecasting performance of the competing models, we conduct conditional forecasts in the reserved part of the sample. The estimated coefficients are held fixed at their in-sample values for the whole ex-post forecast evaluation horizon. The interpretation of the results warrants some caution however. Forecasting errors may either be attributed to the instability of the estimated parameters or to true misalignments of the REER with re-

[33]Panel results of Villavicencio (2006) with regard to the the coefficient of *nfa* are very similar. He considers two different panel estimators (DOLS and the pooled mean group estimator), but only according to one (PMGE) *nfa* enters significantly into a long-run relationship with *reer*. Additionally, the estimated coefficients are also very small (around 0.06) and very similar to our estimates.

spect to its estimated equilibrium value. But we believe that the objective of the BEER approach to match the *observed* real exchange rate behavior as closely as possible, should also remain the objective out-of-sample. We therefore regard a model as superior to another if it provides a better out-of-sample fit.[34] One criterion to assess the relative predictive performance of two competing models (within a specific sample) is to compare their root mean squared errors (RMSE), which are calculated as:

$$\text{RMSE} = \sqrt{\frac{\sum (F_t - A_t)^2}{n}}$$

where F_t is the predicted value of the REER in period t, A_t its actual value, and n the number of forecasts. The lower the RMSE, the better is the predictive performance of the model.

Table 8 shows the relative RMSE of different pairs of models. We compare the RMSE from the fully-country specific model delivering the *best* fit in the post-estimation period with the RMSE which are obtained from panel specifications 9 and 10.

As the results obtained from the two panel specifications are similar, we focus our further remarks on the parsimonious specification *M9*. According to the relative RMSE, panel specification *M9* delivers a better forecasting performance for 8 out of 11 countries, the only exceptions being BEL, CHN and JPN. To test whether the forecasting performance is significantly better, we conduct Diebold Mariano tests, which test the null hypothesis of equal expected forecast accuracy against the alternative of different forecasting ability across models. As only conditional forecasts are conducted, the classical Diebold Mariano statistic can be used and no correction for possible autocorrelation among the residuals has to be made since there are no overlapping forecasts. The panel delivers significantly better forecasts for 8 countries (see table 9), the fully country-specific models are significantly better in just two cases (BEL and JPN). This is especially remarkable, since we have given the heterogeneous estimator an 'unfair' advantage by choosing the country-specific model with the best out-of-sample fit prior to comparing it to the performance of the two panel specifications.

Based on these results, we think that pooling the data provides more robust estimates of the impact of underlying fundamentals on the (observed) real exchange rate, although the poolability hypothesis is statistically rejected. Given the better performance of the pooled estimator in terms of RMSE performance, we recommend its use when calculating equilibrium ex-

[34]Considering the limited size of the reserved sample, we certainly cannot be sure, whether this result is robust or simply due to the presence of extraordinary shocks.

Table 8: Relative RMSE: Country-Best vs. Panel Specifications

BEL	*BEL10*	*M9*	*M10*	**JPN**	*JPN20*	*M9*	*M10*
BEL10	1.00			*JPN20*	1.00		
M9	1.57	1.00		*M9*	1.14	1.00	
M10	2.16	1.38	1.00	*M10*	0.82	0.72	1.00
CAN	*CAN12*	*M9*	*M10*	**KOR**	*KOR18*	*M9*	*M10*
CAN12	1.00			KOR18	1.00		
M9	0.64	1.00		*M9*	0.47	1.00	
M10	0.55	0.85	1.00	*M10*	0.16	0.79	1.00
CHN	*CHN8*	*M9*	*M10*	**NLD**	*NLD4*	*M9*	*M10*
CHN8	1.00			*NLD4*	1.00		
M9	1.30	1.00		*M9*	0.54	1.00	
M10	6.14	4.72	1.00	*M10*	0.65	1.20	1.00
GER	*GER20*	*M9*	*M10*	**SWE**	*SWE17*	*M9*	*M10*
GER20	1.00			SWE17	1.00		
M9	0.15	1.00		*M9*	0.12	1.00	
M10	0.12	0.80	1.00	*M10*	0.16	1.30	1.00
IRL	*IRL5*	*M9*	*M10*	**US**	*US16*	*M9*	*M10*
IRL5	1.00			*US16*	1.00		
M9	0.60	1.00		*M9*	0.54	1.00	
M10	0.40	0.67	1.00	*M10*	0.58	1.08	1.00
ITA	*ITA11*	*M9*	*M10*				
ITA11	1.00						
M9	0.41	1.00					
M10	0.66	1.61	1.00				

Note: Only the fully country-specific model providing the lowest RMSE of all respective country specifications is considered here. For all other country-specific models listed in table 2.11 the relative RMSE performance of the reported panel specifications consequently would be even better.

Table 9: Diebold Mariano Tests

	Intercept Country-Best vs. *M9*	Intercept Country-Best vs. *M10*
BEL	-0.02^{***}	-0.05^{***}
CAN	0.02^{**}	0.03^{***}
CHN	-0.00	-0.10^{***}
GER	0.14^{***}	0.15^{***}
IRL	0.04^{***}	0.07^{***}
ITA	0.04^{***}	0.02^{***}
JPN	-0.05^{***}	0.01
KOR	0.11^{**}	0.13^{**}
NLD	0.05^{***}	0.04^{**}
SWE	0.11^{***}	0.11^{***}
US	0.02^{**}	0.02^{**}

Note: ***,** and * denote rejection of the null hypothesis of equal forecasting performance at the significance levels of 1, 5, and 10%. A negative sign implies a higher forecast accuracy of the country-by-country model vs. the pooled model.

change rates. The efficiency gains achieved by the increased sample size and regressor variability seem to outweigh the costs of inducing bias by imposing identical coefficients.[35]

8 Real Effective Exchange Rate Misalignments

To derive currency misalignments, we first re-estimate panel specification *M9* over the full sample. Results are similar compared to the ones obtained over the restricted sample. This underscores the stability of the parameter estimates derived from the panel.[36] Figure 2.2 shows the estimated equilibrium real effective exchange rates (only based on the long-run relationship, lags and leads of first differences are not included) and the historical real effective exchange rates for all countries together with the implied percentage overvaluation. Since the net foreign asset to GDP ratio is not included in our final specification, it is hard to argue that the sharp depreciation of the USD in current years is a direct response to the growing US external liabilities. Based on our results in the preceding sections we would rather say that the direct influence of the US net foreign asset position on the real value of the USD is fairly limited.

For ease of exposition figure 1 shows the US REER together with the estimated equilibrium exchange rate (BEER) and the implied percentage misalignment (right figure). We observe that the US *reer* follows the general

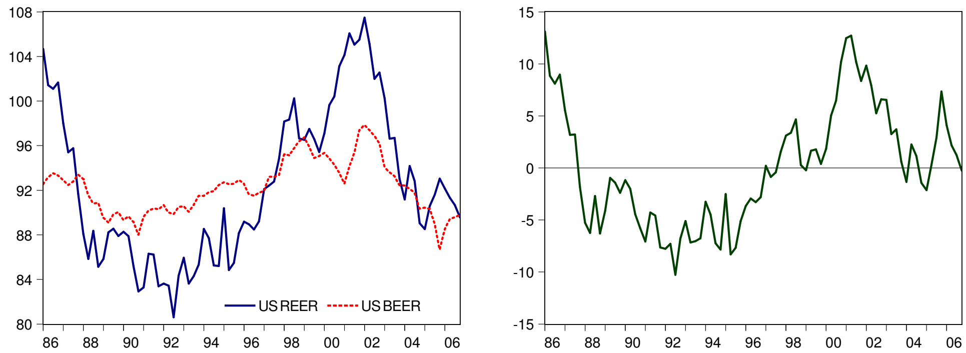


Figure 1: US REER vs. BEER and Misalignment in %

movements of the estimated equilibrium exchange rate. We find the US

[35]See Baltagi (2005). For further details about this method of choosing between pooling or not and another application see Schiavo (2008).

[36]In contrast, re-estimating the fully country-specific models over the full sample sometimes leads to dramatic changes in the size of estimated parameters and even sign changes.

Dollar to be undervalued over a long period from 1987 to 1997 (reaching about 10 percent in early 1993) and to be overvalued from 2000 to 2003 (with a peak overvaluation of about 13 percent in 2001). Since then we observe a correction towards its estimated equilibrium value. At the end of the sample we find the USD to be very close to its predicted value. For Germany we observe an approximate mirror image of this development. Its equilibrium exchange rate is found to be overvalued between 1990 and 1997 and highly undervalued from 2000 to 2003. At the end of 2006 we see a moderate overvaluation of about 3 percent.[37] End of 2006 we find a number of currencies to be misaligned. The Canadian Dollar, the British Pound and the Australian Dollar are 8, 8, respectively 14 percent overvalued in real terms, whereas we see a strong undervaluation of about 22 percent of the Japanese Yen. In contrast to the widespread view that the the Chinese Renminbi is highly undervalued, we find its real value to be in line with the only remaining fundamental.

9 Conclusions

In this paper we estimate equilibrium real effective exchange rates and derive currency misalignments for 15 countries. To account for the observed discretion when it comes to selecting possibly relevant fundamentals or choosing a particular specification, we estimate a large number of specifications for each of the countries individually, and then conduct various pooled estimations. Although the poolability hypothesis is statistically rejected, we find the pooled estimator to perform significantly better in terms of conditional out-of-sample forecasts for most of the countries. This is a remarkable result, since we have given the heterogeneous (country-by-country) estimator an unfair advantage by choosing the country-specific model (of up to 21 possible ones) with the best out-of-sample performance prior to comparing it with two parsimonious panel specifications.

While we find strong evidence in favor of the BS effect, evidence in favor of other commonly hypothesized fundamentals is only weak.

End of 2006 we find two currencies to be significantly overvalued: the Australian and the Canadian Dollar. On the other side, we find the Japanese Yen to be more than 20 percent undervalued. A possible extension to our analysis would be to test the relative forecasting performance of pooled es-

[37]These results are very similar to the ones Bénassy-Quéré et al. (2008) obtain for the 'synthetic' Euro. So the development of the Deutsche Mark seems to be a good proxy for the Euro, although intra-Eurozone trade has not been netted out and therefore trading weights have not been adjusted accordingly.

timators for sub-panels for which the null hypothesis of poolability cannot be rejected. For our panel of 15 countries another 32751 sub-panels consisting of at least two countries could be tested with the help of an iterative procedure. Another option would be to check how 'intermediate' estimators such as shrinkage estimators or the pooled mean group estimator perform out-of-sample in this context.

References

Alberola, E., Cervero, S., Lopez, H. and Ubide, A. (1999) Global equilibrium exchange rates: Euro, Dollar, "ins", "outs", and other major currencies in a panel cointegration framework, IMF Working Paper No. 99/175.

Amano, R. and van Norden, S. (1998) Oil prices and the rise and fall of the US real exchange rate, *Journal of International Money and Finance*, **17**, 299–316.

Baltagi, B. (2005) *Econometric analysis of panel data*, New York: John Wiley & Sons.

Baltagi, B., Bresson, G. and Pirotte, A. (2008) To pool or not to pool?, *Advanced Studies in Theoretical and Applied Econometrics*, **46**, 517.

Baltagi, B. and Griffin, J. (1997) Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline, *Journal of Econometrics*, **77**, 303–327.

Baltagi, B., Griffin, J. and Xiong, W. (2000) To pool or not to pool: Homogeneous versus heterogeneous estimators applied to cigarette demand, *Review of Economics and Statistics*, **82**, 117–126.

Banerjee, A., Marcellino, M. and Osbat, C. (2004) Some cautions on the use of panel methods for integrated series of macroeconomic data, *Econometrics Journal*, **7**, 322–340.

Banerjee, A., Marcellino, M. and Osbat, C. (2005) Testing for PPP: Should we use panel methods?, *Empirical Economics*, **30**, 77–91.

Bayoumi, T., Faruqee, H. and Lee, J. (2005) A fair exchange? Theory and practice of calculating equilibrium real exchange rates, IMF Working Paper No. 05/229.

Bergstrand, J. (1991) Structural determinants of real exchange rates and national price levels: Some empirical evidence, *The American Economic Review*, **81**, 325–334.

Bénassy-Quéré, A., Béreau, S. and Mignon, V. (2008) How robust are estimated equilibrium exchange rates? A panel BEER approach, CEPII Working Paper No. 1, CEPII Working Paper 2008-01.

Breitung, J. and Pesaran, M. (2005) Unit roots and cointegration in panels, Bundesbank Discussion Paper No. 42.

Chinn, M. (1997) Sectoral productivity, government spending and real exchange rates: Empirical evidence for OECD countries, NBER Working Paper No. 6017.

Chinn, M. and Johnston, L. (1996) Real exchange rate levels, productivity and demand shocks: Evidence from a panel of 14 countries, IMF Working Paper No. 97/66.

Clark, P. and MacDonald, R. (1998) Exchange rates and economic fundamentals: A methodological comparison of BEERs and FEERs, IMF Working Paper No. 98/67.

Driscoll, J. and Kraay, A. (1998) Consistent covariance matrix estimation with spatially dependent panel data, *Review of Economics and Statistics*, **80**, 549–560.

Driver, R. and Westaway, P. (2005) Concepts of equilibrium exchange rates, Bank of England Working Paper No. 248.

Dufrenot, G. and Yehoue, E. (2005) Real exchange rate misalignment: A panel co-integration and common factor analysis, IMF Working Paper No. 05/164.

Edwards, S. (1989) *Real exchange rates, devaluation, and adjustment. Exchange rate policy in developing countries*, The MIT Press.

Edwards, S. and Savastano, M. (1999) Exchange rates in emerging economies: What do we know? What do we need to know?, NBER Working Paper No. 7228.

Elbadawi, I. (1994) *Estimating Long-Run Equilibrium Real Exchange Rates*, Peterson Institute.

Elliott, G., Rothenberg, T. and Stock, J. (1996) Efficient tests for an autoregressive unit root, *Econometrica*, **64**, 813–836.

Faruqee, H. (1995) Long-run determinants of the real exchange rate: A stock-flow perspective, *IMF Staff Papers*, **42(1)**, 80–107.

Frenkel, J. and Mussa, M. (1985) Asset markets, exchange rates and the balance of payments, NBER Working Paper No. 1287.

Funke, M. and Rahn, J. (2005) Just how undervalued is the Chinese Renminbi?, *The World Economy*, **28**, 465–489.

Genberg, H. (1978) Purchasing power parity under fixed and flexible exchange rates, *Journal of International Economics*, **8**, 247–276.

Égert, B. (2004) Assessing equilibrium exchange rates in CEE Acceding countries: Can we have DEER with BEER without FEER? A critical survey of the literature, BOFIT Discussion Paper No. 1/2004.

Goldfajn, I. and Valdes, R. O. (1999) The aftermath of appreciations, *Quarterly Journal of Economics*, **114**, 229–262.

Hoechle, D. and Basel, S. (2007) Robust standard errors for panel regressions with cross-sectional dependence, *Stata Journal*, **7**, 281–312.

Hurlin, C. and Mignon, V. (2007) Second Generation Panel Unit Root Tests, HAL Working Paper.

Im, K., Pesaran, M. and Shin, Y. (2003) Testing for unit roots in heterogeneous panels, *Journal of Econometrics*, **115**, 53–74.

Kao, C. (1999) Spurious regression and residual-based tests for cointegration in panel data, *Journal of Econometrics*, **90**, 1–44.

Kao, C. and Chiang, M. (1997) On the estimation and inference of a cointegrated regression in panel data, *Advances in Econometrics*, **15**, 179–222.

Kim, B. and Korhonen, I. (2005) Equilibrium exchange rates in transition countries: Evidence from dynamic heterogeneous panel models, *Economic Systems*, **29**, 144–162.

Lane, P. and Milesi-Ferretti, G. (2007) The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities, 1970-2004, *Journal of International Economics*, **73**, 223–250.

Levin, A., Lin, C. and Chu, C. (2002) Unit root tests in panel data: Asymptotic and finite-sample properties, *Journal of Econometrics*, **108**, 1–24.

MacDonald, R. (1998) What determines real exchange rates? The long and the short of it, *Journal of International Financial Markets, Institutions & Money*, **8**, 117–153.

MacDonald, R. (2000) Concepts to calculate equilibrium exchange rates: An overview, Economic Research Group of the Deutsche Bundesbank Discussion Paper No. 3.

MacDonald, R. and Dias, P. (2007) Behavioural equilibrium exchange rate estimates and implied exchange rate adjustments for ten countries., University of Glasgow Working Paper No. 12.

MacKinnon, J. (1990) Critical values for cointegration tests, University of California Economics Working Paper No. 4.

Maddala, G., Trost, R., Li, H. and Joutz, F. (1997) Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators, *Journal of Business & Economic Statistics*, **15**, 90–100.

Melecký, M. and Komárek, L. (2007) The behavioral equilibrium exchange rate of the Czech Koruna, *Transition Studies Review*, **14**, 105–121.

Ng, S. and Perron, P. (2001) Lag length selection and the construction of unit root tests with good size and power, *Econometrica*, **69**, 1519–1554.

Nilsson, K. (2004) Do fundamentals explain the behaviour of the Swedish real effective exchange rate?, *The Scandinavian Journal of Economics*, **106**, 603–622.

Patterson, K. (2000) *An introduction to applied econometrics: A time series approach*, Macmillan New York.

Pedroni, P. (2004) Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis, *Econometric Theory*, **20**, 597–625.

Persyn, D. and Westerlund, J. (2008) Error-correction–based cointegration tests for panel data, *Stata Journal*, **8**, 232–241.

Pesaran, M. H. (2006) Estimation and inference in large heterogeneous panels with a multifactor error structure, *Econometrica*, **74**, 967–1012.

Pesaran, M. H. (2007) A simple panel unit root test in the presence of cross-section dependence, *Journal of Applied Econometrics*, **22**, 265–312.

Rapach, D. and Wohar, M. (2004) Testing the monetary model of exchange rate determination: A closer look at panels, *Journal of International Money and Finance*, **23**, 867–895.

Ricci, L., Milesi-Ferretti, G. and Lee, J. (2008) Real exchange rates and fundamentals: A cross-country perspective, IMF Working Paper No. 08/13.

Rogoff, K. (1996) The Purchasing Power Parity puzzle, *Journal of Economic Literature*, **34**, 647–668.

Roy, S. (1959) *Some aspects of multivariate analysis*, vol. 1-2, Wiley New York.

Saikkonen, P. (1991) Asymptotically efficient estimation of cointegration regressions, *Econometric Theory*, **7**, 1–21.

Schiavo, S. and Vaona, A. (2008) Poolability and the finance-growth nexus: A cautionary note, *Economic Letters*, **98**, 144–147.

Stein, J. (1995) The natural real exchange rate of the US Dollar and determinants of capital flows, *Fundamental determinants of exchange rates*, pp. 38–84.

Stock, J. and Watson, M. (1993) A simple estimator of cointegrating vectors in higher order integrated systems, *Econometrica*, **61**, 783–820.

Taylor, A. and Taylor, M. (2004) The Purchasing Power Parity Debate, *The Journal of Economic Perspectives*, **18**, 135–158.

Vaona, A. (2008) STATA tip: A quick trick to perform a Roy-Zellner test for poolability in STATA, *Quaderni della facoltà di Scienze economiche dell'Università di Lugano*.

Villavicencio, A. (2006) Real equilibrium exchange rates: A panel data approach for advanced and emerging economies, *Economia Internazionale/International Economics*, **4**, 59–81.

Westerlund, J. (2007) Testing for error correction in panel data, *Oxford Bulletin of Economics and Statistics*, **69**, 709–748.

Williamson, J. (1994a) Estimates of FEERS, *Estimating Equilibrium Exchange Rates*, **177**, 245.

Williamson, J. (1994b) *Estimating equilibrium exchange rates*, Peterson Institute.

Williamson, J. (2009) Exchange rate economics, *Open Economies Review*, **20**, 123–146.

Wren-Lewis, S. and Driver, R. (1998) *Real exchange rates for the year 2000*, Peterson Institute.

Zellner, A. (1962) An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias, *Journal of the American Statistical Association*, **57**, 348–368.

10 Appendix

Table 10: Trade Weights for Panel

	AUS	BEL	CAN	CHN	FRA	GER	IRL	ITA	JPN	KOR	MEX	NLD	ESP	SWE	CHE	UK	US
AUS	0.000	0.022	0.029	0.090	0.044	0.089	0.013	0.047	0.180	0.053	0.011	0.023	0.017	0.018	0.016	0.079	0.269
BEL	0.008	0.000	0.014	0.028	0.164	0.206	0.019	0.088	0.047	0.012	0.008	0.077	0.048	0.023	0.022	0.117	0.121
CAN	0.006	0.008	0.000	0.032	0.020	0.033	0.005	0.016	0.051	0.015	0.031	0.009	0.007	0.005	0.005	0.026	0.732
CHN	0.017	0.015	0.031	0.000	0.041	0.087	0.006	0.035	0.258	0.085	0.016	0.020	0.016	0.012	0.010	0.037	0.313
FRA	0.007	0.064	0.016	0.033	0.000	0.218	0.021	0.120	0.053	0.015	0.009	0.050	0.094	0.020	0.034	0.104	0.142
GER	0.009	0.054	0.017	0.046	0.144	0.000	0.023	0.116	0.073	0.020	0.014	0.066	0.056	0.028	0.056	0.106	0.172
IRL	0.008	0.029	0.016	0.020	0.081	0.140	0.000	0.052	0.071	0.018	0.010	0.039	0.027	0.016	0.024	0.218	0.230
ITA	0.010	0.046	0.016	0.036	0.158	0.231	0.017	0.000	0.055	0.018	0.011	0.050	0.069	0.019	0.040	0.088	0.135
JPN	0.021	0.016	0.032	0.162	0.041	0.086	0.014	0.033	0.000	0.086	0.021	0.022	0.015	0.012	0.015	0.047	0.378
KOR	0.017	0.011	0.025	0.150	0.033	0.065	0.010	0.030	0.239	0.000	0.020	0.017	0.014	0.008	0.009	0.039	0.313
MEX	0.003	0.006	0.041	0.022	0.015	0.036	0.004	0.014	0.045	0.016	0.000	0.007	0.009	0.005	0.005	0.014	0.759
NLD	0.007	0.068	0.014	0.034	0.110	0.223	0.022	0.083	0.060	0.017	0.008	0.000	0.043	0.030	0.026	0.127	0.126
ESP	0.006	0.045	0.012	0.029	0.224	0.199	0.016	0.125	0.044	0.015	0.012	0.046	0.000	0.019	0.021	0.098	0.090
SWE	0.013	0.042	0.019	0.046	0.095	0.201	0.019	0.068	0.069	0.017	0.013	0.062	0.037	0.000	0.024	0.118	0.156
CHE	0.008	0.030	0.014	0.028	0.114	0.286	0.020	0.103	0.063	0.014	0.010	0.040	0.030	0.017	0.000	0.081	0.140
UK	0.013	0.051	0.022	0.033	0.116	0.178	0.061	0.074	0.066	0.020	0.009	0.063	0.046	0.028	0.027	0.000	0.194
US	0.014	0.017	0.196	0.088	0.049	0.090	0.020	0.036	0.168	0.050	0.154	0.020	0.013	0.012	0.014	0.060	0.000

Note: Original trade weights taken from Bayoumi et al. (2005), and rescaled so that weights sum to 1.

Table 11: Country-by-Country DOLS and FMOLS Estimates

		NFA	TB	NTT	GOV	TOT	OPEN	$\bar{R}^2$	α
BEL10	DOLS		0.89	0.83***		0.35		0.75	−0.13
			(0.61)	(0.12)		(0.34)			(0.08)
	FMOLS		0.67	0.83***		0.31		0.72	
			(0.47)	(0.11)		(0.25)			
CAN10	DOLS		−1.05**	1.34***		0.82***		0.96	−0.24***
			(0.52)	(0.13)		(0.24)			(0.14)
	FMOLS		−0.94**	1.35***		0.78***		0.93	
			(0.45)	(0.11)		(0.16)			
CAN12	DOLS		−0.75	1.48***	−0.34***	0.64***		0.97	−0.33*
			(0.46)	(0.12)	(0.12)	(0.22)			(0.19)
	FMOLS		−0.70*	1.48***	−0.31***	0.63***		0.94	
			(0.40)	(0.10)	(0.11)	(0.14)			
CAN14	DOLS		−0.34	1.01***	−0.25*	0.74***	−0.35**	0.97	−0.33**
			(0.42)	(0.26)	(0.14)	(0.20)	(0.16)		(0.20)
	FMOLS		−0.45	0.94***	−0.16	0.71***	−0.36***	0.95	
			(0.34)	(0.18)	(0.10)	(0.12)	(0.10)		
CAN21	DOLS			1.05***	−0.28**	0.72***	−0.37***	0.97	−0.24***
				(0.22)	(0.12)	(0.19)	(0.13)		(0.12)
	FMOLS			1.00***	−0.19*	0.68***	−0.38***	0.95	
				(0.18)	(0.10)	(0.12)	(0.11)		

continued on next page

		NFA	TB	NTT	GOV	TOT	OPEN	$\bar{R}^2$	α
CHN2	DOLS	-0.76^{***}		1.85^{***}	-0.67^{***}			0.79	-0.37^{**}
		(0.27)		(0.24)	(0.12)				(0.12)
	FMOLS	-0.38^{*}		1.63^{***}	-0.49^{***}			0.66	
		(0.21)		(0.23)	(0.09)				
CHN5	DOLS	-0.75^{***}		2.37^{***}	-0.52^{***}	0.87^{***}		0.91	-0.37^{***}
		(0.16)		(0.17)	(0.08)	(0.18)			(0.15)
	FMOLS	-0.31^{*}		1.97^{***}	-0.37^{***}	0.64^{***}		0.70	
		(0.17)		(0.21)	(0.08)	(0.18)			
CHN6	DOLS	-0.47		1.63^{***}	-0.55^{***}		0.00	0.83	-0.47^{***}
		(0.30)		(0.36)	(0.18)		(0.36)		(0.17)
	FMOLS	-0.53^{**}		1.21^{***}	-0.33^{***}		-0.50^{*}	0.73	
		(0.25)		(0.30)	(0.12)		(0.29)		
CHN8	DOLS		0.95	1.71^{***}				0.54	-0.20^{**}
			(0.94)	(0.42)					(0.08)
	FMOLS		0.64	1.61^{***}				0.47	
			(0.85)	(0.40)					
CHN9	DOLS		-0.92	1.12^{***}	-0.43^{***}			0.75	-0.25^{*}
			(0.72)	(0.30)	(0.09)				(0.15)
	FMOLS		-1.17^{*}	1.04^{***}	-0.43^{***}			0.70	
			(0.61)	(0.27)	(0.09)				
CHN10	DOLS		0.15	2.45^{***}		1.39^{***}		0.81	-0.24^{*}
			(0.56)	(0.25)		(0.23)			(0.14)

continued on next page

		NFA	TB	NTT	GOV	TOT	OPEN	$\bar{R}^2$	α
	FMOLS		0.07	2.20***		1.13***		0.62	
			(0.58)	(0.28)		(0.21)			
CHN12	DOLS		−0.54	1.90***	−0.24***	0.95***		0.84	−0.33*
			(0.55)	(0.29)	(0.08)	(0.25)			(0.20)
	FMOLS		−0.98**	1.56***	−0.29***	0.68***		0.76	
			(0.48)	(0.24)	(0.08)	(0.17)			
CHN14	DOLS		−0.90	1.64***	−0.28*	0.82***	0.00	0.88	−0.33*
			(0.64)	(0.41)	(0.14)	(0.27)	(0.28)		(0.20)
	FMOLS		−1.17**	1.27***	−0.26**	0.66***	−0.17	0.77	
			(0.53)	(0.27)	(0.12)	(0.17)	(0.22)		
CHN15	DOLS			1.39***				0.51	−0.18***
				(0.31)					(0.06)
	FMOLS			1.38***				0.49	
				(0.29)					
CHN16	DOLS			1.42***	−0.36***			0.70	−0.28**
				(0.21)	(0.09)				(0.14)
	FMOLS			1.41***	0.36***			0.66	
				(0.20)	(0.08)				
CHN20	DOLS			1.36***	−0.34**		−0.03	0.76	−0.30*
				(0.42)	(0.16)		(0.38)		(0.15)
	FMOLS			1.19***	−0.28**		−0.20	0.68	
				(0.33)	(0.13)		(0.27)		
GER18	DOLS			1.04***			−0.95***	0.85	−0.06

continued on next page

		NFA	TB	NTT	GOV	TOT	OPEN	$\bar{R}^2$	α
				(0.33)			(0.13)		(0.08)
	FMOLS			0.93***			−0.95***	0.83	
				(0.23)			(0.08)		
GER20	DOLS			1.09***	−0.13		−0.96***	0.84	−0.05
				(0.36)	(0.29)		(0.13)		(0.09)
	FMOLS			0.98***	−0.16		−0.97***	0.83	
				(0.24)	(0.23)		(0.10)		
IRL5	DOLS	−0.11***		1.77***	−0.19**	−0.17		0.89	−0.30***
		(0.03)		(0.33)	(0.08)	(0.14)			(0.11)
	FMOLS	−0.09***		1.80***	−0.19***	−0.08		0.80	
		(0.02)		(0.27)	(0.06)	(0.12)			
ITA11	DOLS		−2.44	−1.00			−1.70***	0.86	−0.03
			(0.68)	(1.37)			(0.55)		(0.08)
	FMOLS		−2.49	−0.91			−1.63***	0.82	
			(0.52)	(0.96)			(0.40)		
JPN6	DOLS	−1.12***		1.58**	1.12***		−6.23***	0.92	−0.39***
		(0.22)		(0.62)	(0.20)		(0.67)		(0.14)
	FMOLS	−1.26***		1.72***	1.13***		−5.58***	0.89	
		(0.19)		(0.49)	(0.17)		(0.55)		
JPN7	DOLS	−1.04***		1.64**	1.03***	0.10	−5.92***	0.92	−0.38***
		(0.21)		(0.62)	(0.19)	(0.29)	(0.99)		(0.13)
	FMOLS	−1.21***		1.78***	1.08***	−0.00	−5.70***	0.89	
		(0.18)		(0.47)	(0.17)	(0.17)	(0.71)		

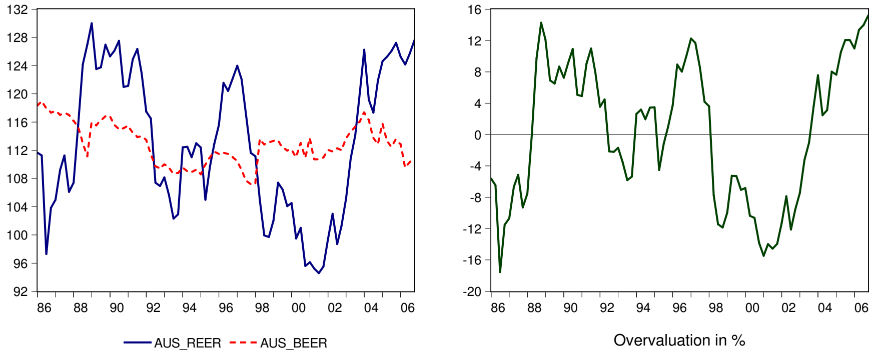
continued on next page

		NFA	TB	NTT	GOV	TOT	OPEN	$\bar{R}^2$	α
JPN20	DOLS			2.19*	0.28		−6.43***	0.76	−0.22*
				(1.12)	(0.26)		(1.08)		(0.11)
	FMOLS			1.88**	0.31		−6.32***	0.73	
				(0.79)	(0.20)		(8.84)		
KOR18	DOLS			0.11			−1.39***	0.76	−0.05
				(0.23)			(0.21)		(0.08)
	FMOLS			0.04			−1.46***	0.67	
				(0.23)			(0.20)		
NLD4	DOLS	−0.20***		0.65***			−0.57***	0.63	−0.20***
		(0.05)		(0.21)			(0.10)		(0.06)
	FMOLS	−0.19***		0.62***			−0.54***	0.59	
		(0.04)		(0.18)			(0.08)		
SWE17	DOLS			0.87***		1.38***		0.93	−0.06
				(0.13)		(0.23)			(0.45)
	FMOLS			0.87***		1.42***		0.87	
				(0.13)		(0.22)			
US1	DOLS	0.05		2.68***				0.70	−0.17***
		(0.21)		(0.55)					(0.04)
	FMOLS	0.03		2.70***				0.64	
		(0.19)		(0.52)					
US8	DOLS		−6.78	0.25				0.83	−0.29***
			(1.37)	(0.63)					(0.08)
	FMOLS		−5.67	0.83**				0.78	

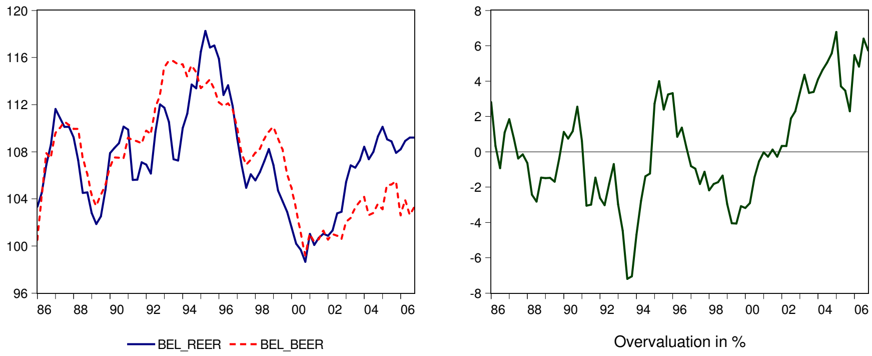
continued on next page

		NFA	TB	NTT	GOV	TOT	OPEN	$\bar{R}^2$	α
			(1.00)	(0.41)					
US9	DOLS		-5.60	0.90	0.19			0.85	-0.22^{**}
			(1.46)	(0.66)	(0.12)				(0.11)
	FMOLS		-4.80	1.22^{***}	0.15			0.80	
			(0.91)	(0.42)	(0.11)				
US15	DOLS			2.78^{***}				0.67	-0.13^{**}
				(0.42)					(0.06)
	FMOLS			2.76^{***}				0.65	
				(0.39)					
US16	DOLS			2.88^{***}	0.25			0.77	-0.09
				(0.45)	(0.15)				(0.07)
	FMOLS			2.85^{***}	0.23			0.66	
				(0.40)	(0.15)				

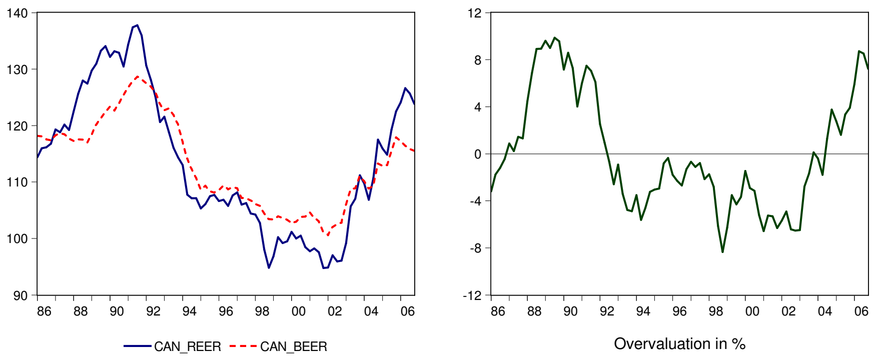
Note: Reported are only specifications, which fulfill the following three conditions: First, all series included in the specification are tested to be I(1) in the *individual* country's unit root tests, secondly, the null hypothesis of no cointegration is rejected at the 5% level according to MacKinnon critical values, and third, *reer* adjusts to deviations from long-run equilibrium. ***,** and * denote the significance levels of 1, 5, and 10%. $\hat{\alpha}$ denotes the estimated adjustment parameter in the error correction form of the model.



(a) AUS

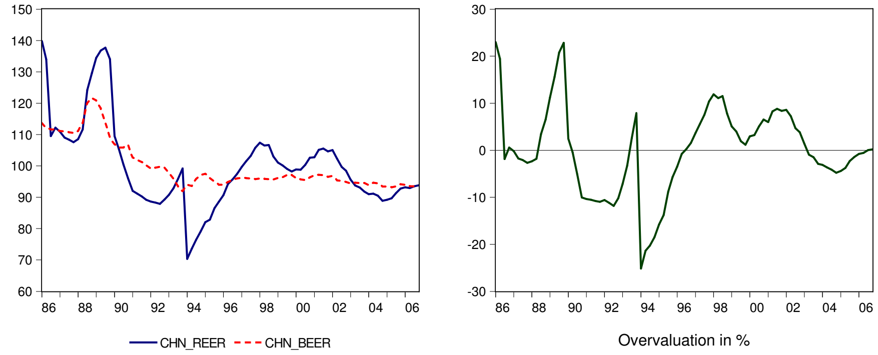


(b) BEL

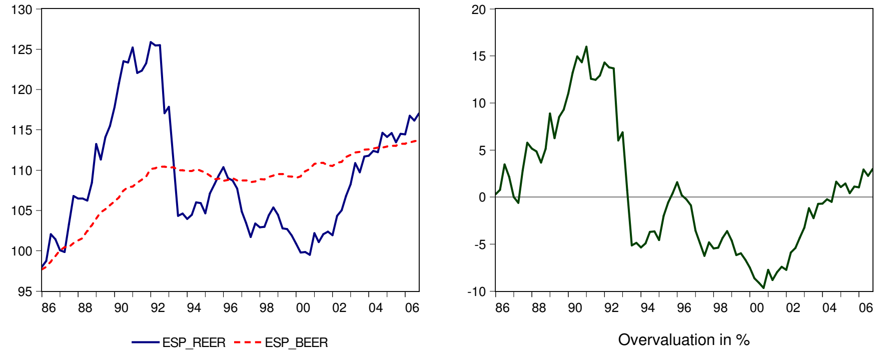


(c) CAN

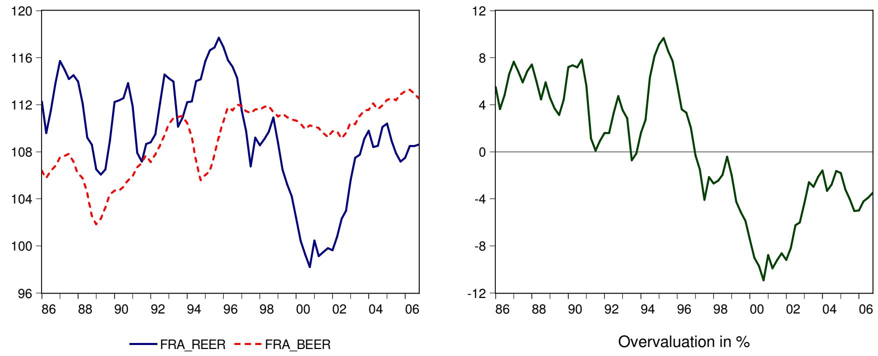
Figure 2: REER vs. BEER and Overvaluation in Percent



(d) CHN

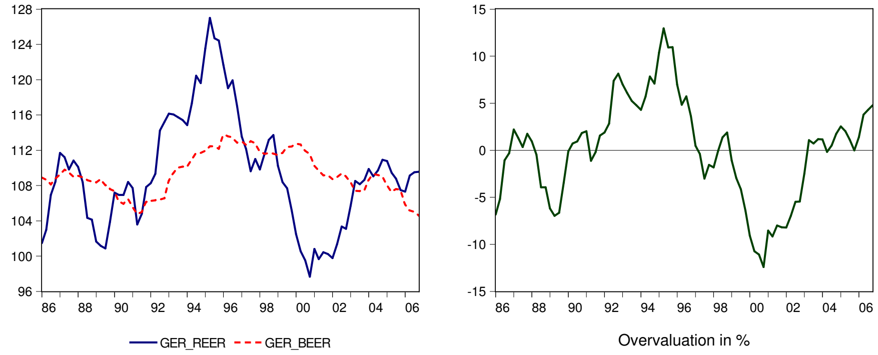


(e) ESP

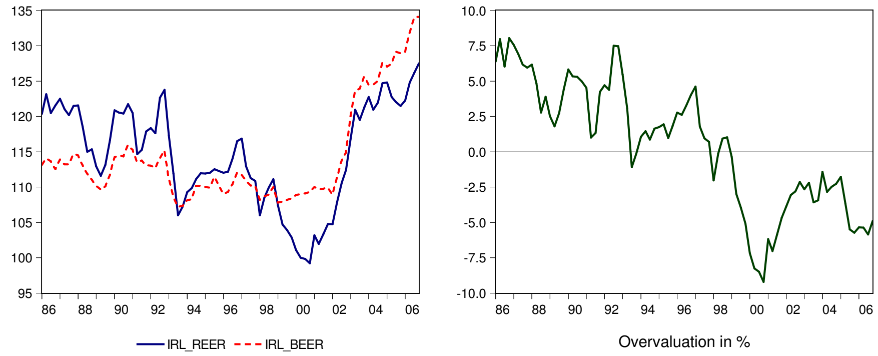


(f) FRA

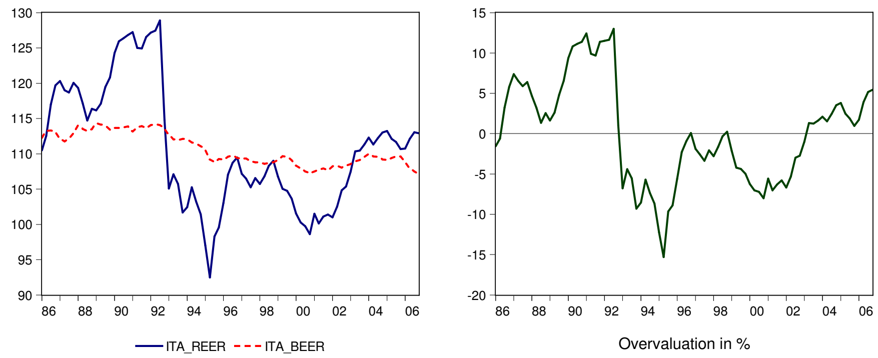
Figure 2: REER vs. BEER and Overvaluation in Percent, contd.



(g) GER

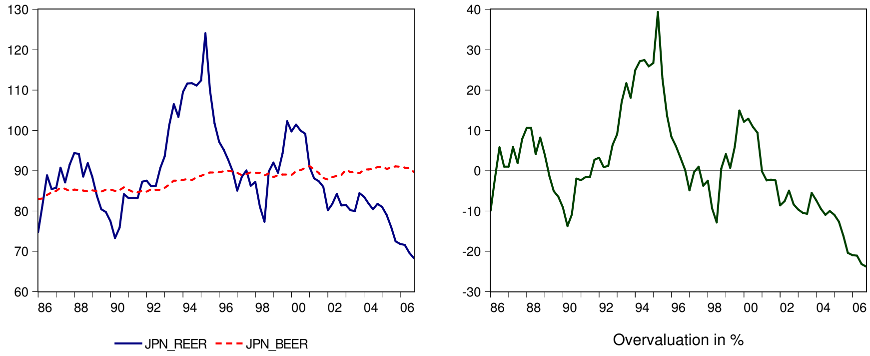


(h) IRL

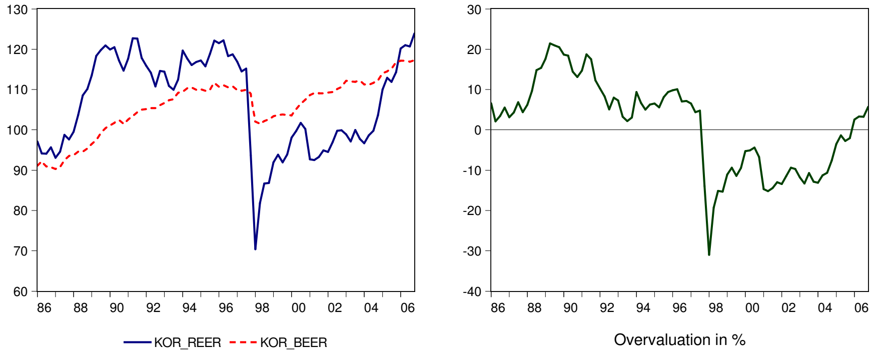


(i) ITA

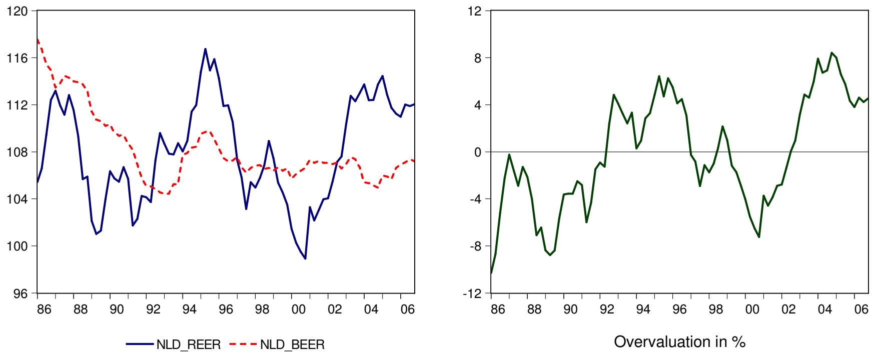
Figure 2: REER vs. BEER and Overvaluation in Percent, contd.



(j) JPN



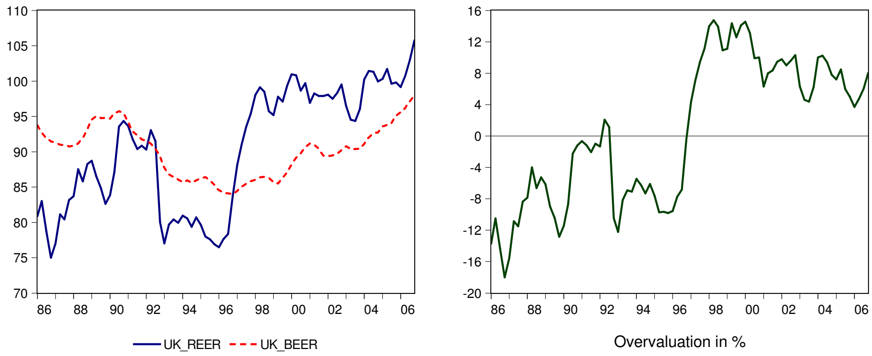
(k) KOR



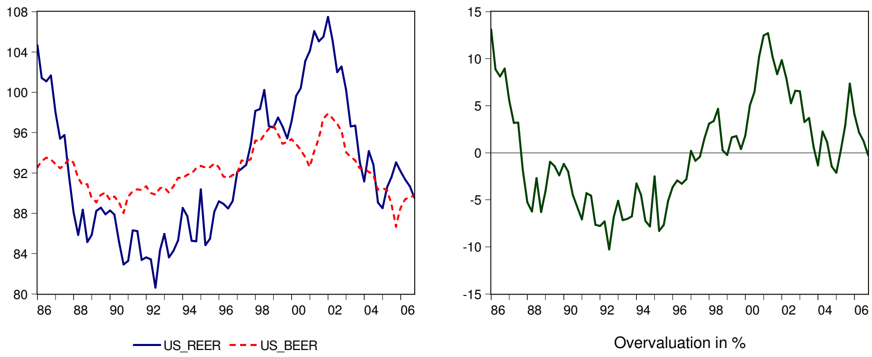
(l) NLD

Figure 2: REER vs. BEER and Overvaluation in Percent, contd.

(m) SWE



(n) UK



(o) US

Figure 2: REER vs. BEER and Overvaluation in Percent, contd.