Herwartz, Helmut; Reimers, Hans-Eggert

Working Paper

Testing the purchasing power parity in pooled systems of error correction models

SFB 373 Discussion Paper, No. 2000,79

Provided in Cooperation with:
Collaborative Research Center 373: Quantification and Simulation of Economic Processes, Humboldt University Berlin

Suggested Citation: Herwartz, Helmut; Reimers, Hans-Eggert (2000) : Testing the purchasing power parity in pooled systems of error correction models, SFB 373 Discussion Paper, No. 2000,79, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin, https://nbn-resolving.de/urn:nbn:de:kobv:11-10048053

This Version is available at:
https://hdl.handle.net/10419/62231

Testing the Purchasing Power Parity in Pooled Systems of Error Correction Models

by

Helmut Herwartz* and Hans-Eggert Reimers

Institute of Statistics and Econometrics
Humboldt–University Berlin
Spandauer Str. 1
D - 10178 Berlin
GERMANY
Tel.: +30 2093-5725

Hochschule Wismar
University of Technique,
Business and Design
Postfach 12 10
D - 23952 Wismar
GERMANY
Tel.: +3841 753601

Abstract

In this paper we test the purchasing power parity for the post Bretton Woods period for 18 main industrial countries. As base currencies we use alternatively the Deutsche mark, the Japanese yen, and the US dollar. We employ error correction models for single countries and on the level of pooled equations allowing efficient inference on domestic and foreign price elasticities of nominal exchange rates. Likelihood ratio tests are applied to infer on linear restrictions implied by the economic relationship. Critical values for these tests are estimated by means of the wild bootstrap that copes with heterogeneous error distributions and contemporaneous correlation within a pooled system of single equations. Furthermore, the tests are performed recursively in order to address the issue of time dependence of our results. We find that the purchasing power parity provides an accurate description of exchange rate dynamics if the Deutsche mark or the Japanese yen are used as base currencies. Specified towards the US dollar we reject the economic model. It turns out that this overall conclusion is not invariant with respect to the investigated sample period.

Keywords: Purchasing power parity, Panel cointegration, Wild bootstrap.
JEL Classification: F31, C22

*The research of this paper was carried out within the Sonderforschungsbereich 373 at the Humboldt-University Berlin and was printed using funds made available by the Deutsche Forschungsgemeinschaft. Helpful comments from Franz Seitz are gratefully acknowledged. Earlier versions of the paper were presented at seminars of the economic departments of the University in Rostock and Hamburg. We thanks the participants for their useful suggestions.

1 Introduction

The purchasing power parity (PPP) postulates a stable long run relationship between the exchange rate of two currencies and the price levels of the involved economies. Stationarity of real exchange rates implies necessary and sufficient conditions for PPP to hold. From these arguments one may follow two particular strategies to test the economic relationship (for comprehensive surveys of the empirical literature on the PPP see, for example, Froot and Rogoff 1995, Rogoff 1996, and Edison et al. 1997). In the first place one may test the unit root hypothesis for real exchange rates. Single country or panel unit root tests are in widespread use to infer on the PPP hypothesis (see e.g. Oh 1996, Lothian 1997, Kuo and Mikkola 1998, Papell and Theodoridis 1998). Panel unit root tests as introduced by Im et al. (1995) or Levin and Lin (1992), however, do not take cross section correlation into account. Feasible GLS procedures are natural candidates to cope with cross sectional dependence (see e.g. O'Connell 1998, Papell 1997). Apart from being a two step procedure the latter device has the disadvantage that critical values of employed test statistics are not pivotal and have to be simulated. The simulation design itself should correspond to the particular investigated data set. Empirically, it turns out that panel unit root testing of real exchange rates typically requires highly parameterized test regressions (see e.g. Papell 1997) which may also affect the empirical performance of this approach. It is worthy to note that in the framework of the Levin and Lin (1992) model and feasible GLS methods specific cross sectional restrictions are a priori assumed to hold in the pooled system.

As a second approach, we employ a dynamic error correction model (ECM) where current changes of the exchange rate are related to lagged violations of the long run relationship. Foregoing contributions to PPP testing by means of a cointegration approach are for instance Cheung and Lai (1993) or Corbae and Ouliaris (1991). On the pooled level Pedroni (1997) investigates the PPP hypothesis for 25 currencies as a link between exchange rates and price ratios employing residual based cointegration tests. Whereas this two step procedure might be preferable in presence of more than one identified cointegrating relation within a dynamic system the error correction approach adopted in this paper is preferable for dynamic systems containing one endogenous variable and only one long-run relationship. Assuming the domestic and foreign price levels to be weakly exogenous for the determination of long run elasticities the conditional single equation ECM allows asymptotically efficient inference by

means of standard OLS procedures (see e.g. Boswijk 1995). In this framework the likelihood principle is applicable to test symmetry and proportionality of foreign and domestic price elasticities separately.

Turning to a pooled level of single equations the prominent likelihood ratio (LR) statistic looses its pivotal property if the underlying error terms exhibit cross sectional correlation. As recently shown by Herwartz and Neumann (2000) critical values for common LR-statistics can be obtained by a particular resampling scheme, namely the wild bootstrap. Initially introduced by Wu (1986) this heteroskedasticity consistent method can be implemented for both single equation models and pooled systems thereof. In the latter case cross sectional dependence can be handled conveniently without any parametric formalization as it is required by GLS methods.

From an economic point of view, the estimation of single dynamic ECM equations allows for different speeds of adjustments to the long run equilibrium. This is expected in view of differential trading costs and heterogeneity of shocks affecting global economies. For instance Flôres et al. (1999) find that the adjustment speed of the Japanese yen against the Deutsche mark is much slower compared to the adjustment dynamics of European currencies against the Deutsche mark. Therefore unconstrained empirical models are preferable to counterparts imposing a priori restrictions. For the econometric analysis of the PPP we find that the conditional ECM provides a parsimonious representation of nominal exchange rate dynamics.

Using quarterly data we test the PPP during the post–Bretton Woods era for 18 currencies of industrial economies. As base currencies we take alternatively the Deutsche mark (DEM), the Japanese yen (JPY), and the US dollar (USD). The data sets include three variables for each relation, namely the exchange rate, the foreign and the domestic price level. We infer on cointegration on the pooled level and test the symmetry and proportionality hypotheses for the long run price elasticities separately. In addition, restrictions on the long run parameters are tested recursively using samples of increasing size.

Our results concerning PPP depend on the employed base currency and to some extent on the investigated sample period. For the entire sample period (1973.1-1998.3) the PPP restrictions cannot be rejected on the pooled level if the ECMs are specified against the DEM or the JPY. Employing the USD as base currency, however, both symmetry and

proportionality hypotheses are rejected on the pooled level. Interestingly we find that the PPP model specified towards the JPY is rejected if the sample period terminates during the so-called bubble economy period. If the sample period ends in the late or early 1980s we cannot reject the PPP model specified towards the USD. During both periods inflationary pressures can be observed for the main industrial economies. In addition, employing the DEM as base currency the support of PPP is considerably reduced during the 1980s. With samples ending around the German unification boom the PPP is rejected within a pooled system of Euro-zone currencies.

The remainder of the paper is organized as follows. In the next section the PPP relation is defined and the implied linear restrictions are given. In section 3 the adopted bootstrap method is briefly provided. Section 4 contains the discussion of empirical results, Section 5 summarizes and concludes.

2 The PPP and implied linear restrictions

The PPP hypothesis states that the price level in two countries, expressed in the same currency, ought to determine the 'fundamental' or 'equilibrium' exchange rate

$$E = \lambda \frac{P}{P^*} \tag{1}$$

where E denotes the exchange rate as the domestic price of one unit of foreign currency, P and P^* denote indices of domestic and foreign prices, and λ is a constant. This relation is analyzed in logarithms, thus

$$e = \ln \lambda + (p - p^*), \tag{2}$$

where small letters denote variables measured in logarithms. In its absolute form, PPP is often rejected empirically because equation (1) does not hold exactly for any pair of countries over any time period. This may be explained by transportation costs, restrictions on trade, imperfect competition on goods markets, differences in calculating the price levels, etc. (see Krugman and Obstfeld 1996, pp. 412-414). PPP may hold in the long-run, however, if deviations from the relationship in (2) are transitory.

It has been recognized that exchange rates and prices are nonstationary. Adopting a time series approach a necessary condition for PPP is that e_t, p_t and p_t^* are cointegrated.

Defining a drift term $\nu = \ln \lambda$ the general long run relation is:

$$e_t = \nu + \beta_1 p_t + \beta_2 p_t^*. \tag{3}$$

The PPP essentially implies two properties which can be tested as linear restrictions within a cointegration framework. First, proportionality of domestic and foreign price elasticities with respect to the exchange rate is conveniently investigated by testing the hypothesis $H_0 : \beta_1 = -\beta_2 = 1$. In practice, price ratios and the exchange rate may not evolve proportionally. Due to transportation costs, measurement errors, differences in price indices or different national productivity shocks the cointegrating parameters may differ from unity (see Froot and Rogoff 1995). Second, symmetry of relative prices and the exchange rate is expressed as $H_0 : \beta_1 = -\beta_2$. Note that the symmetry hypothesis cannot be tested by means of common unit root tests.

A growing body of empirical literature on PPP investigates stationarity of real exchange rates $q_t = e_t - p_t + p_t^*$ within a pooled system. Anker (1999) gives a survey of some studies in this framework. Table 1 displays the most important characteristics of particular contributions. Two main empirical strategies can be distinguished. Firstly, panel OLS approaches are followed in the spirit of Levin and Lin (1992) or Im et al. (1995). Providing a pivotal test statistic if the cross sectional and time series dimension get indefinitely large, both models do not account for cross sectional correlation. In addition, the Levin and Lin (1992) test a priori imposes specific cross equation restrictions. Secondly, feasible GLS methods are applied to take cross sectional correlation into account. Depending on the particular investigation additional regressors are included in the test regressions. Following the latter approach the unit root test statistic is no longer pivotal and one has to simulate critical values, which depend on the cross sectional dimension, the sample size and the postulated underlying data generating process.

Applying panel unit root tests the null hypothesis of non stationary real exchange rates is often rejected (see Table 1), thus supporting the PPP model. An important exception is O'Connell (1998). He shows that the standard practice of calculating all real rates relative to one currency can lead to cross sectional dependence in time series panel data. The presence of contemporaneous correlation within a system may cause standard tests to suffer from size distortions. Anker (1999) points out that O'Connell (1998) restricts the autocorrelation coefficients to be identical. Under this restriction the panel test for a unit root is supposed

to be invariant with respect to the employed base currency.

Table 1 further reveals that the available evidence on PPP depends to some extent on the base currency. Whereas real exchange rates measures towards the DEM are typically found to be stationary the PPP model specified for the USD is more often rejected. Furthermore, with respect to the JPY the validity of PPP is still an open issue. Most investigations mentioned in Table 1 use consumer price indices (CPI). However, the conclusions concerning the validity of the postulated relationship appear to be unaffected if an alternative price variable as, for instance, the wholesale price index (WPI) is used.

Instead of testing on nonstationarity of real exchange rates another approach to test the PPP model is to exploit the cointegration property by specifying a dynamic error correction model (see e.g. Cheung and Lai 1993, Corbae and Ouliaris, 1991). Along such lines Edison et al. (1997) reject the PPP for 13 major industrial countries applying the Johansen (1991) maximum likelihood procedure. Employing the Horvath-Watson (1995) procedure these authors find moderate evidence in favor of PPP during the post–Bretton Woods period. Dynamic vector autoregressive model specifications may suffer from a large dimension of the assumed parameter space. Therefore, single equation approaches may be preferable. The statistical efficiency or even validity of such a device, however, may suffer from ignoring valuable information provided in the multivariate model. As a parsimonious model we are going to test the PPP by means of the conditional ECM. Under specific assumptions this approach can be shown to allow asymptotically efficient estimation and inference.

In this paper LR-tests in single equations as well as in pooled systems are performed. For the former tests the χ^2-distribution may be used to obtain asymptotically valid critical values. On the pooled level, however, the χ^2-distribution does not apply in presence of cross sectional error correlation. In addition, the small sample performance of the second order asymptotic approximation may suffer from model errors which are not independently and identically distributed. The wild bootstrap method provides asymptotically valid critical values even for the case of heteroskedastic error terms and in presence of cross sectional dependence. To cope with these violations of standard modelling assumptions via the advocated bootstrap scheme does not require any parametric specification. Thus, our results are not based on feasible two step procedures.

We investigate the PPP model towards the most important international currencies, namely DEM, JPY and USD. Furthermore the stability of our results is illustrated by means of a recursive implementation of the LR-test.

3 Methodology

3.1 The conditional ECM

Tests of the PPP within a cointegration approach require the specification of a long run relation and additional short run dynamics of the involved variables. Assuming presample values to be available we first consider the single country case, i.e.

$$\Delta e_t = \nu + \alpha(e_{t-1} + \beta_1 p_{t-1} + \beta_2 p^*_{t-1}) + \gamma_1 \Delta p^*_t + \gamma_2 \Delta p_t + \gamma_3 \Delta e_{t-1} + u_t,\ t = 1, \ldots, T. \quad (4)$$

The adopted single equation approach of estimating long run equilibrium relations is asymptotically efficient if the following assumptions hold (see Banerjee et al. 1993, Chapter 6, and Boswijk 1995): Firstly, the involved variables e_t, p_t and p^*_t are assumed to be integrated of order one. Secondly, there exists one and only one linear combination of these variables providing stationary residuals, i.e. e_t, p_t, and p^*_t are cointegrated with cointegrating rank 1. Thirdly, p_t and p^*_t are weakly exogenous for the estimation of β_1 and β_2, and, finally, the error sequence u_t is a martingale difference sequence (mds), i.e. $E[u_t|u_{t-1}, \ldots, u_1] = 0$. In order to ensure the mds property of u_t it might be necessary to augment equation (4) with further (lagged) stationary explanatory variables.

Apart from estimation issues the assumptions provided above are also sufficient to perform asymptotically efficient inference by means of common $t-$ratios and $F-$type statistics on joint significance of selected parameters. For instance, the $t-$ratio of $\hat{\alpha}\,(t_{\hat{\alpha}})$ obtained from OLS-routines is asymptotically normally distributed if the cointegration assumption holds. If the variables are not cointegrated $ec_t = (e_t + \beta_1 p_t + \beta_2 p^*_t)$ is nonstationary. In this case the regression in (4) is "unbalanced" since a nonstationary variable cannot sensibly explain the dynamics of a stationary process. Therefore α should be zero if ec_t is integrated of order one. Thus, a test on significance of α is implicitly also a test of the null hypothesis of cointegration. Under the alternative of no cointegration, however, $t_{\hat{\alpha}}$ does not follow a standard distribution. For this case Kremers et al. (1992) show that $t_{\hat{\alpha}}$ is distributed some-

where between the standard normal distribution and the distribution of the Dickey-Fuller t–statistic.

3.2 The LR-Test

Boswijk (1993) discusses the WALD–statistic for testing composite hypothesis involving the long run parameters β_1 and β_2. Since the WALD–test is not invariant with respect to reformulations of the null hypothesis one may regard a LR-test as a promising alternative device. In the following the LR-test is briefly outlined since this statistic easily allows to perform pooled tests if the model in (4) is just the $n-$th equation in a N-dimensional system of ECMs. Under the null and the alternative hypothesis the regression model provided in (4) may be given compactly as:

$$y = X^0\delta^0 + u^0 \quad \text{and} \quad y = X^1\delta^1 + u^1. \tag{5}$$

In (5) $y = (\Delta e_1, \Delta e_2, \ldots, \Delta e_T)'$, and X^0 and X^1 are design matrices which contain the sets of explanatory variables under the null and alternative hypothesis, respectively. Similarly δ^0 and δ^1 denote vectors of parameters representing the two hypotheses. Employing X^0 and X^1 alternatively to provide OLS-estimates of the sum of residual squared errors, RSS^0 and RSS^1, the LR-statistic is immediately obtained as:

$$\mathrm{LR}_n = T \ln\left(\frac{\mathrm{RSS}^0}{\mathrm{RSS}^1}\right). \tag{6}$$

Under the null hypothesis LR_n is $\chi^2(q)$-distributed where q is the number of restrictions imposed under the null hypothesis (see e.g. Lütkepohl 1991, Chapter 12). Herwartz and Neumann (2000) illustrate that the mds assumption made for u_t is sufficient to obtain an asymptotic χ^2–distribution if the hypothesis of interest involves only long run parameters. If the LR-test is applied to test coefficients of stationary variables, for instance α or $\gamma_j, j = 1, 2, 3$, an asymptotic χ^2–distribution of the LR-statistic is only obtained if the error terms are independent and identically distributed. In particular, homoskedasticity of u_t is required in this case. Accounting for heteroskedasticity, however, may even be sensible for tests of long run parameters in order to improve the small sample properties of the LR-test. Two scenarios of heteroskedasticity are regarded to be of specific importance for the present investigation. On the one hand, one may conjecture that the error terms in (4) have different unconditional variances in the beginning of the sample compared to its final period in the

sequel of changing macroeconomic policies. On the other hand, since the seminal article by Engle (1982) there is little doubt about the finding that variables measured on financial markets as e.g. exchange rates show patterns of conditional heteroskedasticity.

Now consider a set of empirical models as in (4), i.e.

$$\Delta e_{nt} = \nu_n + \alpha_n(e_{nt-1} + \beta_1 p_{t-1} + \beta_{2n} p^*_{nt-1}) + \gamma_{1n}\Delta p^*_{nt} + \gamma_2 \Delta p_t + \gamma_{3n}\Delta e_{nt-1} + u_{nt}, \quad (7)$$

$t = 1, \ldots, T$, $n = 1, \ldots, N$. We are interested in testing a specific null hypothesis to hold in the pooled system. Assuming the error terms of the N equations to be contemporaneously uncorrelated a convenient generalization of the statistic given in (6) is:

$$\begin{aligned} \mathrm{LR}_p &= \sum_{n=1}^{N} \mathrm{LR}_n \\ &= T \sum_{n=1}^{N} \ln\left(\frac{RSS_n^0}{RSS_n^1}\right). \end{aligned} \quad (8)$$

The LR-statistic given above is provided in Lütkepohl (1991), Chapter 12, in a related context. By construction of the test the error terms are allowed to be heteroskedastic across equations. Furthermore, it is implicitly assumed in (8) that T observations are available for each equation. If this assumption is violated, however, the LR_p-statistic can easily be modified. Under the null hypothesis LR_p is $\chi^2(qN)$–distributed. Note that the likelihood approach outlined here is particularly convenient for the case of a finite (small) cross section dimension N.

3.3 The wild bootstrap

In the realistic case of contemporaneous correlation across equations the test statistic given in (8) looses its asymptotic $\chi^2(qN)$-distribution. In presence of only weak cross correlation this distribution may still provide reasonable critical values. Strictly speaking, however, the LR_p statistic can only be regarded as a descriptive tool in this case. Herwartz and Neumann (2000) show that a particular bootstrap procedure, namely the wild bootstrap, can be implemented to obtain critical values for the LR_p-statistic if error terms u_{nt} are contemporaneously correlated across equations. The wild bootstrap applied in this paper has been initially introduced by Wu (1986) to cope with time varying second order moments of regression error terms. For $F-$type statistics in parametric regression models with random explanatory variables Mammen (1993) proved the validity of wild–bootstrap inference.

Adopting a nonparametric framework, Neumann and Kreiss (1998) show that the convenience of regression type bootstrap procedures is often maintained for autoregressive models if the error terms are martingale difference sequences.

Herwartz and Neumann (2000) provide a detailed investigation of small sample properties of bootstrap inference via the LR_p-statistic. It turns out that the wild bootstrap provides appropriate critical values and reduces considerably size distortions involved by applying critical values from the $\chi^2(qN)$-distribution in small samples. The empirical size improvements are shown to increase with N. In the light of the empirical performance of the LR-statistic Herwartz and Neumann (2000) recommend to use bootstrap procedures to mimic the distribution of LR_n or LR_p in practice. The empirical performance of the LR_p-statistic for stationary time series models is investigated in Herwartz (1998). In such a framework it turns out that the distribution of LR_p is strongly affected by heteroskedastic error terms. Since the reader may not be that familiar with the wild bootstrap we briefly sketch such a resampling scheme in the following for completeness:

1. A matrix representation of the multi equation model (7) is straightforward to obtain. In this case the design matrices X_0 and X_1 are blockdiagonal, y and u contain all observations and error terms for $n = 1, \ldots, N$ in stacked form. OLS estimation of (4) ($N = 1$), or (7) provides test statistics LR_n ($N = 1$) or LR_p and estimates $\hat{\delta}^0$ and $\hat{u}^0$.

2. The empirical model specified under the null hypothesis is used to generate a wild bootstrap sample as follows:

$$\tilde{y} = X^0\hat{\delta}^0 + \tilde{u}.$$

   Single elements $\tilde{u}_{nt}$ of $\tilde{u}$ are obtained from OLS residuals $\hat{u}^0_{nt}$ by mimicking their low order moments and cross correlation properties along the following lines (see e.g. Mammen 1993): Generate a sample of T standard normal and independent random variables v_t and w_t and construct $z_t = v_t/\sqrt{2} + (w_t^2 - 1)/2$. It is immediate to verify that $E[z_t] = 0$, $E[z_t^2] = 1$ and $E[z_t^3] = 1$. Bootstrap error terms are then defined as

$$\tilde{u}_{nt} = z_t\hat{u}^0_{nt}. \tag{9}$$

   The innovations given in (9) match the low order moments of $\hat{u}^0_{nt}$, i.e. $E[\tilde{u}_{nt}] = 0, E[\tilde{u}^2_{nt}] = (\hat{u}^0_{nt})^2, E[\tilde{u}^3_{nt}] = (\hat{u}^0_{nt})^3$. Furthermore cross correlation properties of $\hat{u}^0_{nt}$ are mirrored, since $E[\tilde{u}_{nt}\tilde{u}_{mt}] = \hat{u}^0_{nt}\hat{u}^0_{mt}$, $n, m = 1, \ldots, N$, $n \neq m$.

3. From the generated variables $\tilde{y}$ and the initial design matrices, X^0 and X^1 it is again feasible to compute the LR-statistic, denoted as LR_n^* $(N = 1)$ or LR_p^*.

4. Steps (2) and (3) are performed R times, where R is sufficiently large. For each bootstrap sample the statistic of interest is recorded. For the present analysis we used $R = 1000$ replications.

5. The hypothesis $H_0 : \delta = \delta^0$ is rejected with significance level α if LR_n (LR_p) exceeds the $(1 - \alpha)$-quantile of LR_n^* (LR_p^*).

As a particular advantage of the wild bootstrap it is worthwhile to mention that it is not necessary to resample the regressor series in (4) or (7). Thus it is not necessary to make specific assumptions on the data generating processes of these variables. The test statistics are resampled by means of the same set of conditioning variables throughout. In contrast to feasible (two step) GLS methods the application of the wild bootstrap does not require a parametric formalization of heteroskedasticity or cross sectional correlation. The advocated resampling scheme even allows for time varying patterns of cross sectional dependence.

4 Empirical results

4.1 Data and sample period

We examine the PPP with currencies of 18 industrial countries for the post–Bretton Woods period. The exchange rates are prices of the foreign currency in DEM. Since the Austrian central bank started in 1981 to peg the shilling against the DEM we exclude this currency from the sample.

The used national price level is the deflator of final demand. This deflator contains private and public consumption, investment and exports. It is broadly defined putting a high weight on tradable goods. The data are from the database of the Deutsche Bundesbank and cover the period from the first quarter 1973 to the third quarter 1998. All variables are measured in natural logarithms. To test the PPP hypothesis for alternative base currencies the exchange rates are rebased using the triangular equality, for instance $(E_{\mathrm{YEN/USD}} = E_{\mathrm{DEM/USD}}/E_{\mathrm{DEM/YEN}})$.

The post–Bretton Woods period covers the world wide acceleration of inflation rates in respond to the first oil price shock in 1973 (see e.g. Krugman and Obstfeld 1996, Chapter

19). Due to the existing floating rate regimes country specific monetary policy actions took place. Major central banks adopted a monetary targeting policy to stabilize inflation rates. The less tight monetary policy in the United States seems to weaken the US-Dollar against currencies of other major industrial countries from 1976 to 1979.

In 1979 the European Monetary System (EMS) was founded. It was a formal system of mutually pegged exchange rates. Through a mixture of policy cooperation and realignments the EMS survived. Those EMS members who enforced fluctuation margins belonged to the EMS exchange rate mechanism (ERM). One aim of the EMS was to import the German Bundesbank's credibility as an inflation fighter and to reduce inflation to the low German rate. This resulted in an asymmetric system where Germany set the system's monetary policy while the other countries pegged their exchange rates against the DEM, thus exchange rate targeting was conducted.

The beginning of the EMS almost coincided with the second oil price shock in 1979/80 that induced a world wide recession. Monetary growth was restricted in most industrial countries to limit inflation. In the following a disinflation process started. The USD appreciated against most currencies from 1982 to 1985. Trying to manage exchange rates the Group of Five announced in September 1985 to intervene jointly on the foreign exchange market to depreciate the USD (Plaza announcement). In February 1987 they renewed their effort to cooperate on the foreign exchange market trying to stabilize nominal exchange rates (Louvre accord).

Towards the end of the 1980s inflationary pressures reappeared in the main industrial countries. Inflation was more the result of national developments rather than a global shock. In the 1990s a world wide decline of inflation rates is observed. Moreover in the European Union the convergence of inflation rates becomes apparent. After the Treaty of Maastricht (1992) most members of the European Union succeeded in obtaining inflation rates slightly higher than the most price stable countries.

Summarizing, the post–Bretton Woods era is characterized by different inflation rate phases and exchange rate policies.

4.2 Unit root tests

Results from Augmented Dickey Fuller tests applied to the series of nominal exchange rates and price levels are given in Table 2. The test regressions are alternatively specified with intercept term and both intercept term and deterministic trend. The favored specification is determined applying the BIC criterion. Along similar lines the BIC criterion is used to select an appropriate lag order of the test regression. For all alternative test orders, $k = 0, \ldots, 5$, the same presample period is used.

With respect to the exchange rate series similar conclusions can be drawn for almost all countries. For first differences of exchange rates the unit root hypothesis is rejected at the 1% significance level for all countries. Since for only 4 countries the unit root hypothesis is rejected for level series we regard exchange rate series to be integrated of order 1.

Turning to the price level series our results are ambiguous. The unit root hypothesis is rejected for the majority of countries contained in our sample. With respect to first differences of price levels 9 (of 18) series are found to be nonstationary. As an explanation of the observed ambiguity one may regard structural breaks operating in the price generating processes. A close visual inspection of these series over the sample period reveals that for most countries the price level exhibits varying growth patterns. Allowing structural breaks within the ADF–regressions would again support the assumption that the price level series are integrated of order one for the vast majority of national price levels.

4.3 Model diagnostics and cointegration tests

To investigate PPP equation (4) is estimated for 17 exchange rates using the DEM, the JPY and the USD as base currencies, respectively. Estimating single equation models does not involve any cross section restriction of the error correction coefficient or of parameters governing short run dynamics. For the ease of comparison and for the purposes of the panel approach that follows below the number of fitted lags is equal across single country models. Table 3 (DEM), Table 4 (JPY), and Table 5 (USD) display estimates of the standard deviation of the underlying error terms and a number of diagnostic statistics. In particular, these diagnostics are the Ljung-Box-statistic obtained from 20 estimated residual autocorrelations, F-type tests against homoskedasticity and structural invariance and the ARCH-LM statistic testing the homoskedastic model against an ARCH model of order 1. The F-type statis-

tics are computed from OLS regressions applied to two separate subsamples each containing approximately one half of the available observations.

First consider empirical models specified towards the DEM as base currency (see Table 3). From the Ljung-Box statistic minor indications of misspecification of single equation models are obtained. Estimated error terms exhibit significant autocorrelation in four cases at the 5% significance level. Since we find significant heteroskedasticity in almost all models the applied $\chi^2(16)$-distribution is hardly the true asymptotic distribution of the Ljung-Box statistic. We conjecture this statistic to be oversized in the present application. With respect to structural stability of the specified empirical model it turns out that for two equations the hypothesis of structural invariance is rejected at the 10% significance level. Since we investigate 17 data generating processes we do not take this result into account any further. Turning to the assumption of homoskedasticity things look different. Eight processes show error variances in the first part of the sample which differ from their counterparts in the second part at the 5% significance level. For 6 equations the assumption of homoskedasticity is rejected at the 5% level against an ARCH(1) alternative. In summary, heteroskedasticity within single equations cannot be rejected for almost all dynamic models under study.

Taking the JPY (the USD) as base currency the results (see Tables 4 and 5) point in a similar direction as discussed for the DEM. For both sets of estimated ECMs the Ljung-Box-statistics indicate residual autocorrelation for five of 17 equations at the 5% significance level. Evidence of deterministic heteroskedasticity is found in two equations at the 5% level whereas ARCH-effects are diagnosed in 3 cases. Structural instability is only indicated in one case employing the USD as base currency.

We refrain from providing detailed parameter estimates for the three systems of empirical models to economize on space. As mentioned, the estimated loading coefficients are particularly important to test the hypothesis of cointegration. It turns out that all estimated coefficients are negative as it is expected if the error correction mechanism provides an adequate representation of the data generating model. To test the hypothesis $H_0 : \alpha \geq 0$ against $H_1 : \alpha < 0$ on the pooled level one may regard empirical averages $\bar{\alpha} = \frac{1}{N}\sum_{n=1}^{N} \hat{\alpha}_n$ to provide a convenient test statistic. We obtain $\bar{\alpha} = -0.1331$, -0.1347, -0.1081, and $\sigma_{\bar{\alpha}} =$ 1.43E-02, 1.08E-02, 1.09E-02 employing the DEM, JPY, USD as base currency, respectively. Using the standard errors $\sigma_{\bar{\alpha}}$ of the mean estimates $\bar{\alpha}$ to provide a $t-$statistic

we reject $H_0 : \alpha \geq 0$ for the pooled system at any reasonable significance level. Almost the same result would be obtained if the estimated variances of $\hat{\alpha}_n$ are used to compute $\sigma_{\bar{\alpha}}$. Thus we conclude that e_t, p_t and p_t^* are cointegrated at the pooled level.

In contrast to these results Pedroni (1997) finds weaker evidence of cointegration by means of residual based cointegration tests for panel data. In this contribution the long run link between exchange rates and CPI ratios is analyzed for 25 countries in the period June 1973 to December 1994. Apart from major industrial economies the sample includes additionally Turkey, New Zealand, South Africa, Chile, Mexico, India, Korea, and Pakistan. Our results are more in line with Jacobson and Nessen (1998). They analyze PPP between Germany, Great Britain, Japan, and the US over the period 1930 - 1996 using multivariate cointegration techniques. Within a seven dimensional system of three exchange rates and four corresponding price variables Jacobson and Nessen (1998) find a cointegration rank of three as it can be expected under PPP.

As mentioned, the average values of $\bar{\alpha}$ are between -0.108 and -0.135. This implies that on average between 10.8 and 13.5 percent of the deviations from the long run relationship are reduced in one quarter. This value is relatively high compared to results given in Rogoff (1996). He summarizes that PPP deviations tend to damp out at the slower rate of 15 percent per annum. The difference may be explained by the used price variables. The price deflator of final demand is more volatile compared to the CPI used by Rogoff (1996).

4.4 Testing linear restrictions

Sufficient conditions for PPP are tested by imposing linear restrictions on the cointegrating vector. The results of the tests for the symmetry ($H_0 : \beta_1 = -\beta_2$) and the proportionality hypothesis ($H_0 : \beta_1 = -\beta_2 = 1$) are also given in Table 3 to Table 5. We provide the LR-statistics (LR_n) and the corresponding p-values obtained from the asymptotic $\chi^2(q)$–distribution on the one hand and from wild bootstrap inference on the other.

Using the DEM as base currency (Table 3) and taking critical values from a χ^2-distribution the symmetry (proportionality) hypothesis is rejected at the 5% significance level in 3 (6) single country equations. Resampling the LR-statistics by means of the wild bootstrap we obtain marginal significance levels for both hypotheses which are in almost all cases larger than the corresponding p-values obtained from the second order asymptotic approximation.

At the 5% significance level symmetry or proportionality is rejected for two single country models.

Different test results with respect to the competing inference procedures are also obtained for the remaining base currencies. Turning to the JPY the symmetry (proportionality) hypothesis is rejected at the 5% significance level in 2 (3) cases applying critical values from the χ^2-distribution. The number of rejections reduces if the wild bootstrap scheme is used to provide critical values. In this case we do not reject the symmetry hypothesis anymore. In addition, we find only one violation of proportionality which is significant at the 5% level.

Summarizing our results based on the DEM or the JPY we find strong evidence for the PPP hypothesis. The support of the economic relationship is considerably weaker if the USD serves as base currency. Regarding the marginal significance levels estimated by the wild bootstrap procedure we reject the symmetry (proportionality) hypothesis for 4 (6) single equation models at the 5% level.

Both linear restrictions are also tested for the pooled system with the results shown in Table 6. For the symmetry and proportionality hypothesis we obtain test statistics of 36.8 and 79.52 (DEM), 21.2 and 58.7 (JPY), and 60.48 and 85.65 (USD), respectively. Comparing these statistics with critical values obtained from a χ^2-distribution with 17 and 34 degrees of freedom we conclude that only the symmetry hypothesis cannot be rejected if the system is based on the JPY. All remaining test statistics indicate rejection of the respective hypotheses at any reasonable significance level. From the discussion in Section 3 we know that the χ^2-distribution does not deliver valid critical values on the pooled level in presence of cross section error correlation. The bootstrap approach provides appropriate critical values even in this case. It turns out, that for those systems based on the DEM or the JPY the estimated marginal significance levels increase to at least 15%, thus providing strong evidence in favor of the PPP within the pooled system. Even for the pooled system based on the USD the estimated marginal significance levels are higher compared to the χ^2-distribution. It turns out, however, that the symmetry (proportionality) hypothesis still has to be rejected at the 1% (5%) significance level if the distributions of the LR-statistics are mimicked by the wild bootstrap.

Employing the DEM as base currency it is worth to test the PPP hypotheses for two separately pooled systems, namely the Euro zone and the non-Euro zone. As mentioned

above the EMS was an asymmetric system. The DEM was the monetary anchor and the other members conducted exchange rate targeting to obtain inflation convergence. Most of the non-Euro zone currencies float freely against the DEM. Within both pooled systems linear restrictions implied by the PPP are rejected at the 5% level if critical values are taken from the χ^2-distribution. Both restrictions, however, cannot be rejected at conventional significance levels if critical values are generated by the wild bootstrap method.

The presented results imply that the PPP hypothesis seems to be valid in the long run for currencies of the main industrial countries for the post–Bretton Woods period measured against the DEM and JPY. With respect to the former base currency our results also apply for two particular subsystems, namely the Euro zone and the non-Euro zone. This implies that PPP indicates the direction and extent of exchange rate dynamics in a system of pegged and floating exchange rates. The PPP may also be an important indicator of future adjustments of exchange rates in the exchange rate mechanisms II of the European Union. Taking the USD as base currency we are led to reject the PPP on the pooled level.

4.5 Stability of the test results

So far our results on testing the PPP hypothesis are based on the analysis of the entire available sample information. Papell and Theodoridis (1998) point out that the evidence of long-run PPP is actually stronger than it was a decade ago. One the one hand such a finding may be due to an improved performance of statistical procedures that can be expected from an analysis of additional data. On the other hand, economic factors may contribute to the emerging concensus in favor of PPP. Similar to Papell and Theodoridis (1998) we conduct recursive tests of restrictions implied by the economic relationship. If the actual evidence in favor of the PPP hypotheses can be addressed to purely statistical effects we would expect a smooth pattern of LR-type statistics, or equivalently, of the associated marginal significance levels. More volatile patterns of the latter quantity would indicate periods of stronger and weaker support of the PPP model. To implement the recursive testing scheme we start with a sample covering the period 1973, first quarter to 1978, first quarter. Adding a quarter of observations the tests are recursively performed up to 1998, third quarter. Note that such a recursive scheme is particularly useful if the underlying data generating process is stable. From the single equation diagnostics given in Table 3 to Table 5 we do not obtain almost

any indication of structural invariance.

Figure 1 displays recursive LR-test results with the DEM as base currency. As before critical values for these statistics are obtained from the χ^2-distribution and the wild bootstrap procedure. We distinguish three pooled systems, namely the Euro-zone (medium panels), the non-Euro zone (upper panels), and the aggregate of both (top panels). Left (right) hand side panels display results for tests on symmetry (proportionality). Similarly, but only for one aggregate system, Figure 2 shows recursive results for empirical models specified with the JPY (upper panels) and the USD (lower panels) as base currency.

Except for one pooled model and hypothesis test (JPY as base currency, $H_0 : \beta_1 = -\beta_2$) critical values estimated via the bootstrap procedure yield marginal significance levels which differ considerably from those obtained under a χ^2-distribution. Since the former model is supposed to have better size properties in small samples and accounts conveniently for cross sectional correlation we discuss the empirical results only for this approach.

Taking the DEM as base currency and regarding the entirely pooled system the proportionality hypothesis cannot be rejected at the 5% level throughout the period 1979, first quarter to 1998, third quarter. We observe two rejections of the symmetry hypothesis in 1995. In general, marginal significance levels turn out to be somewhat more volatile for tests of the symmetry hypothesis compared to the proportionality assumption. Similarly, taking the DEM as base currency, marginal significance levels vary more within the system of Euro-zone currencies relative to the pool of remaining countries. Within these subsystems we obtain two particular periods where proportionality towards the DEM is rejected at the 5% level. Within the Euro-zone the marginal significance levels of proportionality tests decrease during a period starting in 1982 and terminating with the German unification. For this subsystem we reject proportionality at the 5% level around the German unification in 1990. Note that inflation rates in the Euro zone differed considerably in the 1980s. Due to the unification boom in Germany the domestic prices increase. Furthermore, in 1992 a realignment of the exchange rates took place. With respect to the subsystem of non-Euro zone countries the proportionality hypothesis is rejected for a two year period covering 1995. In 1995 a strong appreciation of the DEM against the USD and the currencies of the EMS took place. The dollar depreciation seems to be an implication of the Mexico crisis in the end of 1994 and the depreciation of European currencies might be explained by the suspicion

against the European Monetary Union. The DEM depreciates from 1996 to 1998. During this period the evidence in favor of PPP increases for both LR-tests and all pooled systems specified with the DEM as base currency.

Taking the JPY as base currency we cannot reject symmetry at the 10%-level for almost all recursive samples. An analogous result is not obtained for the proportionality hypothesis. This assumption is rejected at the 5% level if the sample period ends between 1984 and 1989. Extending the sample covering information of the post asset bubble period the proportionality hypothesis cannot be rejected at least with 5% significance.

Finally, consider recursive results for the USD as base currency. We obtain only a few samples supporting the PPP hypothesis. Interestingly a short period where the proportionality hypotheses cannot be rejected almost coincides with those samples where the proportionality towards the DEM is rejected in the pooled system of Euro-zone countries.

In sum, it is apparent that the evidence in favor of the PPP hypothesis depends on the employed base currency and on the sample period. Confirming the results presented by Papell and Theodoridis (1998) the evidence of PPP is uniformly stronger using the DEM as base currency compared to the USD. Specified towards the JPY the PPP cannot be rejected if the relevant some period does not end during the asset bubble economy. In addition, recursive marginal significance levels are not monotone in time. This indicates that the growing evidence in favor of the PPP model is not purely a statistical artefact.

5 Conclusions

The necessary and sufficient conditions of PPP are tested for currencies of major industrial countries against the DEM, the JPY and the USD for the post–Bretton Woods era. We adopt a single equation error correction framework which is convenient to test efficiently on cointegration and against linear restrictions implied by the PPP hypothesis. For all employed base currencies the average of estimated error correction coefficients is significantly less than zero, thus supporting the hypothesis of cointegration on the pooled level.

Linear restritions of long-run elasticities are tested by means of LR-tests. Critical values for the test statistics are obtained alternatively from an asymptotic χ^2-distribution and a bootstrap scheme. The latter device is supposed to have superior size properties in small

samples of error terms which have the mds property but may still exhibit a time varying distribution. In addition, the bootstrap approach copes with cross sectional error correlation on the pooled level. In general, applying a second order asymptotic approximation the the corresponding marginal significance levels are smaller compared to their counterparts estimated via the bootstrap model. On the pooled level, for instance, we reject almost any implication of the PPP model employing critical values of the χ^2-distribution irrespective of the employed base currency. With respect to the bootstrap approach PPP implications are found to be more in line with the sample information at least if the dynamic equations are specified with the DEM or JPY as the base currency. The applied resampling scheme allows for heteroskedastic error terms and cross sectional correlation without requiring any parametric specification thereof.

Furthermore, our results are not invariant with respect to the employed sampling period. Recursive estimates of marginal significance levels of the LR-tests are not monotone in time, thus indicating that the actually emerging consensus in favor of the PPP is not merely a statistical artefact. In particular, we reject the PPP model specified against the JPY for samples ending during the second half of the 1980s, the asset bubble period. Specified towards the DEM the economic relationship is rejected within the pooled system of Euro currencies if the sample period ends around the German unification boom. Interestingly, the empirical model with the USD as base currency is in line with the PPP if the employed sample period terminates around 1990 or during the early 1980s. Both periods are charakterized by inflationary pressures emerging in most industrial economies.

References

Jacobson, T., M. Nessen (1998): *Long-Run Perspectives on Purchasing Power Parity*, discussion paper of Sveriges Riksbank presented at the ESEM98 in Berlin.

Anker, P. (1999): Pitfalls in Panel Tests of Purchasing Power Parity, *Weltwirtschaftliches Archiv*, 135, pp. 437–453.

Banerjee, A., J.D. Dolado, J.W. Galbraith, D.F. Hendry, (1993): *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press, Oxford.

Boswijk, H.P. (1993): On the Formulation of Wald Tests on Long Run Parameters, *Oxford Bulletin of Economics and Statistics*, 55, 137–144.

Boswijk, H.P. (1995). Efficient Inference on Cointegration Parameters in Structural Error Correction Models, *Journal of Econometrics*, 69, 133–158.

Cheung, Y.-W., K.S. Lai (1993): Long-Run Purchasing Power Parity during the Recent Float, *Journal of International Economics*, 34, 181–192.

Corbae, D., S. Ouliaris (1991): A Test of Long-Run PPP Allowing for Structural Breaks, *Economic Record*, 67, 26–33.

Edison, H.J., J.E. Gagnon, W.R. Melick (1997): Understanding the Empirical Literature on Purchasing Power Parity: the Post Bretton Woods Era, *Journal of International Money and Finance*, 16, 1-17.

Engle, R.F. (1982): Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, 50, 987–1008.

Flôres, R., P. Jorion, P.-Y. Preumont, A. Szafarz (1999): Multivariate Unit Root Tests of the PPP Hypothesis, *Journal of Empirical Finance*, 6, 335–353.

Frankel, J.A., A.K. Rose (1996): A Panel Project on Purchasing Power Parity: Mean Reversion within and between Countries, *Journal of International Economics*, 40, 209–224.

Froot, K.A., K. Rogoff (1995): Perspectives on PPP and Long-Run Real Exchange Rates, in *Handbook of International Economics*, vol. III, G. Grossman and K. Rogoff (eds.), chapter 32, 1647-1788.

Herwartz, H. (1998): Testing Periodicity in Time Series Models – A Recommendation of Bootstrap Methods, *Computational Statistics*, 13, 283–300.

Herwartz, H., M. Neumann (2000): Bootstrap Inference in Single Equation Error Correction Models, *mimeo.*

Horvath, M.T.K., M.W.Watson (1995): Testing for Cointegration when some of the cointegrating vectors are known, *Econometric Theory*, 11, 984–1014.

Im, K.S., M.H. Pesaran, Y. Shin (1995): Testing for Unit Roots in Heterogeneous Panels, *University of Cambridge, DAE Working Paper No. 9526.*

Jacobson, T., M. Nessen (1998): Long-Run Perspectives on Purchasing Power Parity, *discussion paper Sveriges Riksbank*, presented at the ESEM98 in Berlin.

Johansen, S. (1991) Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, vol. 59, pp. 1551-1580.

Jorion, P., R.J. Sweeney (1996): Mean Reversion in Real Exchange Rates: Evidence and Implications for Forecasting, *Journal of International Money and Finance*, 15, 535–550.

Kremers, J.J.M., N.R. Ericsson, J.J. Dolado (1992): The Power of Cointegration Tests, *Oxford Bulletin of Economics and Statistics*, 54, 325–348.

Krugman, P., M. Obstfeld (1996): *International Economics, Theory and Policy, 4th edition* Addison-Wesley, Reading.

Kuo, B.-S., A. Mikkola (1998): How Sure Are We About PPP? Panel Evidence with the Null of Stationary Real Exchange Rates, *Discussion paper*, presented at the ESEM98 and EEA98, Berlin.

Levin, A., C.-F. Lin (1992): Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *University of California, USCD Working Paper, 92/23, San Diego.*

Lothian, J.R. (1997): Multi-Country Evidence on the Behavior of Purchasing Power Parity under the Current Float, *Journal of International Money and Finance*, 16, 16–35.

Lütkepohl, H. (1991): *Introduction to Multiple Time Series Analysis*, Springer Verlag, Berlin.

Mammen, E. (1993): Bootstrap and Wild Bootstrap for High Dimensional Linear Models, *The Annals of Statistics*, 21, 255–285.

Meier, C.P. (1997): Assessing Convergence in Purchasing Power Parity: A Panel for 10 OECD Countries, *Weltwirtschaftliches Archiv*, 133, 297–312.

Neumann, M.H., J.-P. Kreiss (1998): Regression-type Inference in Nonparametric Autoregression, *Annals of Statistics*, 26, 1570–1613.

O'Connell (1998): The Overvaluation of Purchasing Power Parity, *Journal of International Economics*, 44, 1–19.

Oh, K.-Y. (1996): Purchasing Power Parity and Unit Root Tests Using Panel Data, *Journal of International Money and Finance*, 15, 405–418.

Papell, D.H. (1997): Searching for Stationarity: Purchasing Power Parity under the Current Float, *Journal of International Economics*, 43, 313–332.

Papell, D.H., H. Theodoridis (1998): Increasing Evidence of Purchasing Power Parity over the Current Float, *Journal of International Money and Finance*, 17, 41–50.

Pedroni, P. (1997): Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests With an Application to the PPP Hypothesis, *Indiana University, Discussion paper.*

Rogoff, K. (1996): The Purchasing Power Parity Puzzle, *Journal of Economic Literature*, 34, 647-668.

Wu, C.F.J. (1986): Jackknife, Bootstrap, and Other Resampling Methods in Regression Analysis (with discussion), *Annals of Statistics*, 14, 1261–1295.

Wu, Y. (1996): Are Real Exchange Rates Nonstationary? Evidence from a Panel-Data Test, *Journal of Money, Credit, and Banking*, 28, 54–63.

	Estimation	Critical values	Data	Base	Unit root
Oh (1996)	Panel OLS, lags	Simulation serial correlations	Price levels from PWT, annual 1960-89	dollar	rejected
Frankel/Rose (1996)	Panel OLS lags	Levin/Lin (1992)	CPI, annual 1948-92	dollar	not rejected
Wu (1996)	Panel OLS, lags, time-specific aggr. effects	Simulation	CPI and WPI, annual, monthly, quarterly 1974:1-93:4	dollar	rejected
Jorion / Sweeney (1996)	GLS	Simulation cross country correlation	CPI, monthly 1973:1-93:12	dollar, mark	rejected
Meier (1997)	GLS, lags	Simulation	Tradables prices, annual 1973-1994	dollar, mark	rejected
Papell (1997)	GLS lags	Simulation serial correlation	CPI quarterly monthly 1973:1-94:4	dollar, mark	dollar: mixed mark: rejected
O'Connell (1998)	GLS lags (restrictions)	Simulation, cross country and serial correlation	CPI, quarterly 1973:2-95:4	depends on sub-sample	not rejected
Lothian (1997)	Panel OLS lags	Simulation serial correlation	CPI annual 1974-90	dollar	rejected
Flôres et al. (1999)	GLS lags	Simulation cross country correlation	CPI monthly 1973:1-94:12	dollar mark	both rejected
Anker (1999)	GLS lags	Simulation serial and cross country correlation	CPI, WPI quarterly 1977:2-97:1	dollar mark	both rejected

Table 1: Panel unit root tests with real exchange rates. Note: In the second column, 'lags' indicates allowance for additional lags in the regressions. In the 'critical values' column it is indicated which kind of correlation under the null is regarded in the simulation.

	Level series						First Differences					
	p_t			e_t			Δp_t			Δe_t		
	ADF	TR	k	ADF	TR	k	ADF	TR	k	ADF	TR	k
	Level Series											
BE	-2.614^c	0	2	-1.467	0	1	-3.297	1	1	-6.341^a	0	0
CA	-4.917^a	0	1	-1.498	0	1	-6.910^a	1	0	-6.771^a	0	0
CH	-1.208	0	2	-3.754^a	0	0	-3.028^b	0	1	-7.792^a	0	0
DK	-1.856	0	3	-2.374	0	1	-1.626	0	2	-6.972^a	0	0
FL	-5.417^a	0	2	-1.437	0	1	-1.684	0	2	-6.824^a	0	0
FR	-2.980^b	0	1	-1.778	0	4	-2.157	0	0	-4.430^a	0	3
GE	-5.518^a	0	0	0.000	0	0	-8.146^a	1	0	0.000	0	0
GR	-1.441	0	4	-1.461	0	0	-0.835	0	3	-8.700^a	0	0
IR	-3.656^a	0	4	-3.813^a	0	0	-2.610	1	3	-8.957^a	1	0
IT	-3.362^b	0	1	-4.347^a	0	0	-3.216^b	0	0	-7.272^a	0	0
JA	-5.051^a	0	4	-1.801	0	1	-4.016^a	0	5	-6.827^a	0	0
NL	-3.027^b	0	1	-1.899	0	4	-4.409^a	0	0	-10.35^a	0	1
NO	-3.038^b	0	2	-1.013	0	1	-4.098^a	1	1	-6.431^a	0	0
PO	-1.912	0	3	-2.763^c	0	1	-0.729	0	3	-5.969^a	1	0
SE	-3.637^b	0	2	-1.174	0	1	-4.610^a	1	1	-7.349^a	0	0
SP	-2.416	0	3	-2.253	0	0	-0.754	0	3	-7.629^a	0	0
UK	-4.883^a	0	1	-2.149	0	1	-2.415	0	1	-7.576^a	0	0
US	-2.580	0	1	-1.629	0	1	-3.844^b	1	0	-7.122^a	0	0

Table 2: Augmented Dickey Fuller Tests for level series and first differences. Selection of lag order $k = 0, .., 5$ order selection and of deterministic terms by means of BIC criterion. Critical values from own simulations using 10000 replications of a random walk with N(0,1) innovations. TR = 1 indicates a deterministic trend within the ADF-regression. a, b, c indicate significance at the 1%, 5%, and 10% level respectively. The following industrial economies are considered: Belgium (BE), Canada (CA), Switzerland (CH), Denmark (DK), Finland (FL), France (FR), Germany (GE), Greece (GR), Ireland (IR), Italy (IT), Japan (JA), Netherlands (NL), Norway (NO), Portugal (PO), Sweden (SE), Spain (SP), United Kingdom (UK), and United States (US).

	$\hat{\sigma}$	Q(20)	het.	struc.	ARCH	LR1	p1	pb1	LR2	p2	pb2
BE	.012	16.5	7.23^a	0.66	11.6^a	2.77	.096	.221	3.56	.169	.471
CA	.047	12.9	1.56^c	1.00	0.20	2.57	.109	.151	3.04	.219	.279
CH	.022	30.5^b	2.02^b	1.31	1.28	6.82	.009	.029	10.1	.006	.015
DK	.013	14.2	2.92^a	0.88	2.36	0.66	.418	.299	0.66	.719	.699
FL	.029	18.7	1.36	0.71	4.15^b	1.10	.294	.294	6.53	.038	.071
FR	.020	56.4^a	8.35^a	2.42^c	15.2^a	0.91	.341	.284	1.00	.606	.727
GR	.035	14.4	6.50^a	1.03	0.69	0.46	.498	.734	0.60	.742	.831
IR	.025	23.8^c	1.65^c	2.57^c	2.92^c	1.83	.176	.209	12.8	.002	.046
IT	.072	14.7	20.4^a	0.41	17.1^a	0.90	.343	.300	7.24	.027	.065
JA	.044	14.8	1.11	0.45	2.61	4.18	.041	.121	5.26	.072	.188
NL	.008	28.5^b	81.2^a	0.35	44.6^a	6.82	.009	.262	10.1	.006	.225
NO	.022	22.6	1.71^b	0.59	9.96^a	0.06	.813	.845	5.30	.070	.137
PO	.025	22.0	5.48^c	0.21	2.37	0.03	.859	.847	0.06	.969	.982
SE	.033	17.9	1.41	1.17	0.87	1.78	.183	.356	2.35	.309	.442
SP	.033	26.9^b	3.82^c	0.51	0.00	3.61	.058	.044	6.44	.040	.137
UK	.039	13.4	1.54^c	0.78	0.68	2.19	.139	.133	4.25	.119	.203
US	.048	9.72	1.27	1.05	0.94	0.13	.719	.736	0.16	.924	.930

Table 3: Estimates and diagnostics for single equation ECMs. Base currency is the Deutsche Mark. Critical values for the Ljung-Box statistic $Q(20)$ are taken from the $\chi^2(16)$-distribution. het: F-type statistic against deterministic heteroskedasticity struc: F-type statistic on parameter stability. Both F-statistics are obtained from residuals of OLS regressions applied to two separate subsamples. ARCH: Lagrange-Multiplier Test of order one against conditional heteroskedasticity. a, b, c indicate significance at the 1%, 5%, and 10% level. LR-test results for the tests of the symmetry hypothesis (LR1) and proportionality hypothesis (LR2). p1 and p2 are p-values obtained from asymptotic approximations, pb1 and pb2 are p-values generated by bootstrap inference.

	$\hat{\sigma}$	Q(20)	het.	struc.	ARCH	LR1	p1	pb1	LR2	p2	pb2
BE	.044	15.5	1.18	1.10	1.24	2.31	.129	.203	6.98	.031	.064
CA	.048	25.1^{c}	1.36	1.58	0.01	0.19	.663	.571	0.20	.906	.894
CH	.043	9.75	1.22	0.32	1.02	2.44	.118	.221	5.30	.071	.158
DK	.044	17.8	1.51^{c}	0.60	3.24^{c}	0.55	.460	.523	2.96	.228	.299
FL	.048	24.9^{c}	1.66^{b}	0.81	0.55	0.00	.987	.986	0.68	.711	.675
FR	.042	27.1^{b}	1.46	0.64	0.66	0.95	.330	.353	10.5	.005	.021
GE	.044	14.8	1.11	0.45	2.61	4.18	.041	.121	5.26	.072	.188
GR	.052	25.9^{c}	1.43	1.13	0.03	1.34	.248	.200	2.03	.363	.341
IR	.045	17.9	1.53^{c}	1.73	4.16^{b}	0.46	.496	.348	0.49	.785	.735
IT	.080	27.0^{b}	3.45^{a}	0.31	12.5^{a}	2.07	.150	.222	6.46	.040	.095
NL	.045	18.5	1.05	0.56	2.22	4.80	.028	.081	4.82	.090	.151
NO	.046	18.4	1.60^{c}	1.06	1.90	0.21	.645	.634	0.96	.620	.656
PO	.046	25.1^{c}	1.00	0.27	3.39^{c}	0.17	.684	.697	4.34	.114	.210
SE	.051	27.2^{b}	1.46	0.77	4.53^{b}	0.07	.786	.786	5.51	.063	.126
SP	.052	31.1^{b}	1.10	0.26	0.17	0.01	.930	.909	0.34	.845	.760
UK	.050	20.2	1.05	0.73	0.02	0.81	.369	.295	1.22	.543	.539
US	.048	28.1^{b}	1.09	1.73	1.36	0.65	.421	.314	0.66	.720	.676

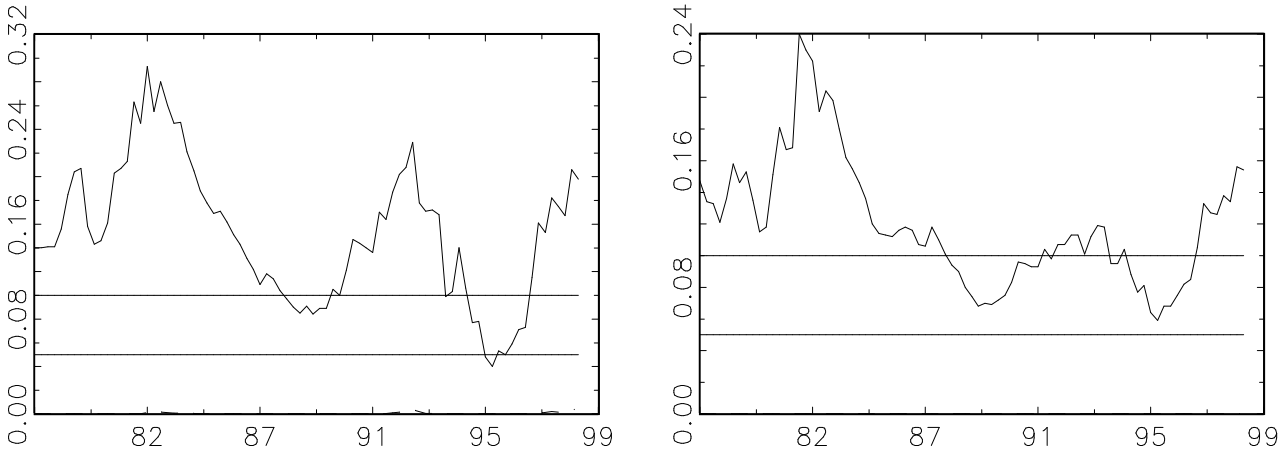
Table 4: Estimates and diagnostics for single equation ECMs. Base currency is the Japanese Yen. See also Table 3.

	$\hat{\sigma}$	Q(20)	het.	struc.	ARCH	LR1	p1	pb1	LR2	p2	pb2
BE	.042	19.3	1.12	0.95	0.10	16.3	.000	.005	16.8	.000	.006
CA	.016	43.3^a	1.15	3.12^b	1.66	2.94	.086	.144	4.06	.131	.175
CH	.052	10.8	1.05	1.71	0.19	2.53	.112	.130	2.86	.239	.267
DK	.043	14.1	1.12	0.94	0.04	1.35	.245	.229	8.74	.013	.013
FL	.045	17.9	1.84^b	0.99	6.22^b	1.03	.311	.272	1.37	.504	.472
FR	.041	12.9	1.33	0.58	1.46	0.14	.708	.700	3.23	.199	.194
GE	.048	9.72	1.27	1.05	0.94	0.13	.719	.736	0.16	.924	.930
GR	.041	40.5^a	1.06	1.52	0.02	5.84	.016	.043	7.84	.020	.029
IR	.041	20.1	1.44	0.85	0.19	0.08	.778	.765	1.35	.509	.517
IT	.067	27.0^b	3.26^a	0.29	7.72^a	0.21	.643	.686	1.64	.440	.548
JA	.048	28.1^b	1.09	1.73	1.36	0.65	.421	.314	0.66	.720	.676
NL	.043	12.5	1.05	0.58	1.51	10.5	.001	.009	11.5	.003	.010
NO	.040	25.3^c	1.63^c	1.54	1.34	0.03	.869	.858	0.03	.984	.987
PO	.043	14.3	1.08	0.82	0.58	4.64	.031	.052	4.65	.098	.129
SE	.045	23.3	1.20	0.72	13.0^a	8.77	.003	.011	9.38	.009	.013
SP	.045	19.6	1.24	1.32	0.28	3.48	.062	.099	3.56	.169	.129
UK	.045	31.2^b	1.07	1.22	2.52	1.92	.166	.146	7.74	.021	.035

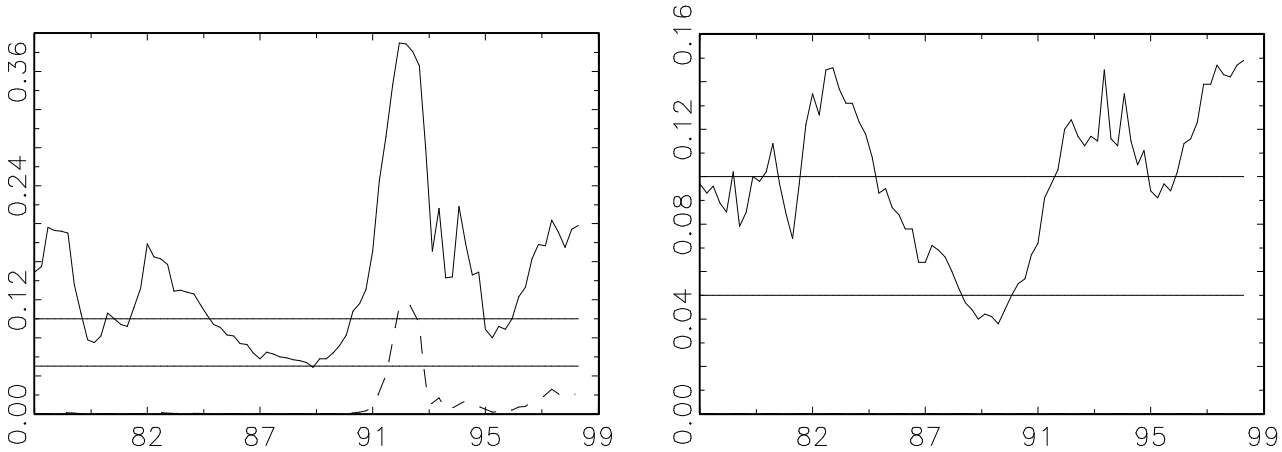
Table 5: Estimates and diagnostics for single equation ECMs. Base currency is the US Dollar. See also Table 3.

			GE		JA	US
		total	Euro-Zone	Non-Euro Z.		
Symmetry	LR_p	36.80	17.96	18.84	21.19	60.48
hypothesis	p1	0.004	0.022	0.027	0.218	0.000
$H_0 : \beta_1 = -\beta_2$	pb1	0.198	0.198	0.228	0.255	0.004
Proportionality	LR_p	79.52	47.84	31.68	58.69	85.65
hypothesis	p2	0.000	0.000	0.024	0.000	0.000
$H_0 : \beta_1 = -\beta_2 = 1$	pb2	0.154	0.149	0.219	0.150	0.017

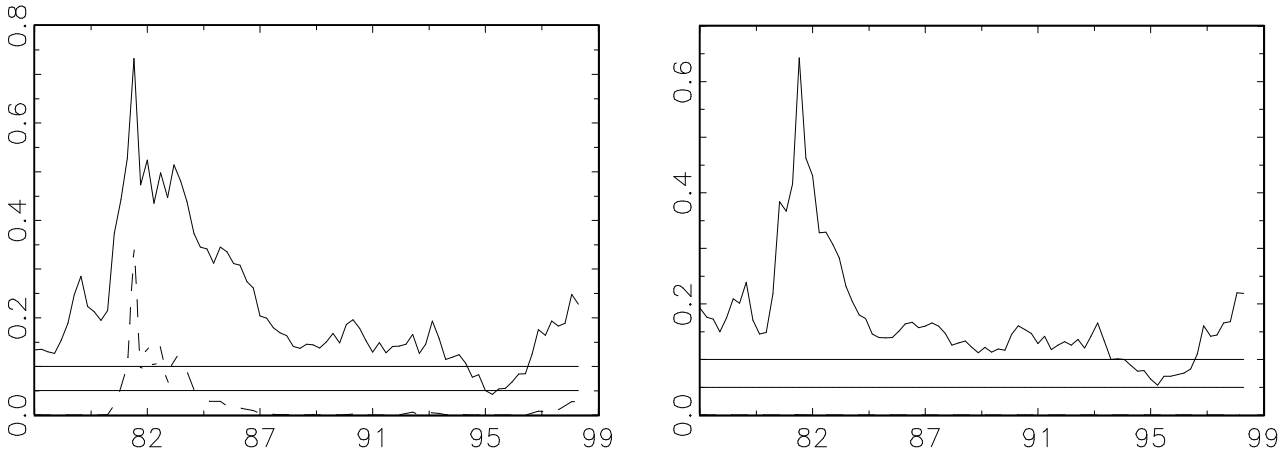
Table 6: Pooled LR-tests on symmetry and proportionality hypotheses using alternative base currencies. With respect to the German Mark LR-test results are also given for two subsystems, namely the Euro Zone and the Non-Euro Zone. The Euro-zone includes the countries BE, FL, FR, IR, IT, NL, PO, SP. The non-Euro-zone contains CH, DK, GR, JA, CA, NO, SE, UK, US. p1 and p2 (pb1 and pb2) denote marginal significance levels obtained from the asymptotic approximation (bootstrap inference).



All countries



Euro zone countries



Non-Euro zone countries

Figure 1: Marginal significance levels for recursive LR-tests for symmetry (left hand side panels) and proportionality (right hand side panels) hypothesis. Base currency is the DEM. Sampling period is 1978.1-1998.3. Dotted line χ^2 critical values, solid line bootstrap inference.

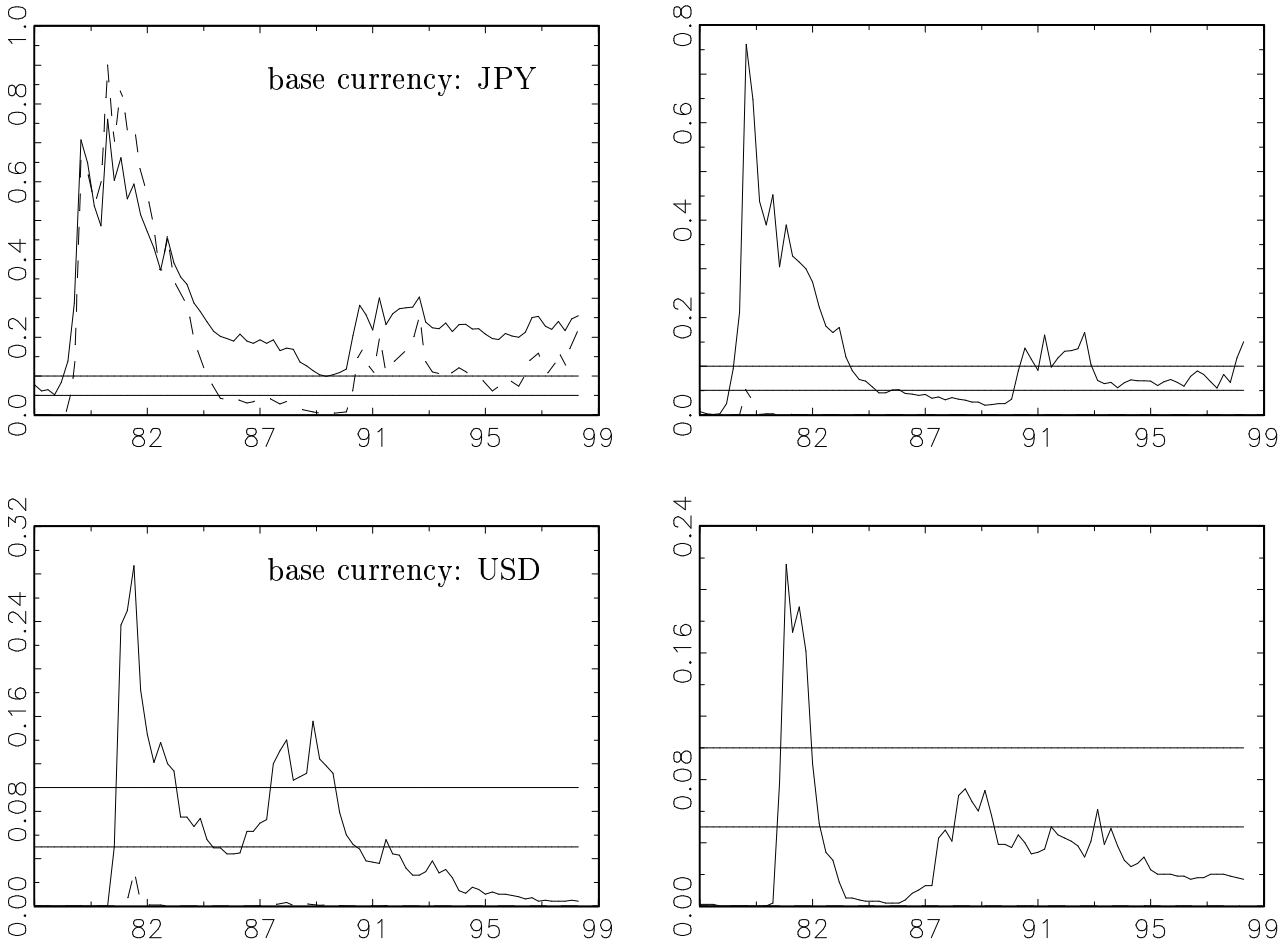


Figure 2: Marginal significance levels for recursive LR-tests for symmetry (left hand side panels) and proportionality (right hand side panels) hypothesis on the level of 17 pooled equations. Base currencies are the JPY (upper panels) and the USD (lower panels). Sampling period is 1978.1-1998.3. Dotted line χ^2 critical values, solid line bootstrap inference.