Purchasing Power Parity:

Evidence from a New Test

by Franc Klaassen [*]

CentER and Department of Economics
Tilburg University

January 25, 1999

Abstract

Most economists intuitively consider purchasing power parity (PPP) to be true. Nevertheless, quite surprisingly, the empirical literature is not very supportive for PPP. In this paper, however, we find evidence in favor of PPP using a new test. The test is embedded in a Markov regime-switching model for the exchange rate, because earlier papers have shown that this model describes the data better than the popular random walk. We allow for PPP by making the regime-switching probabilities depend on the PPP deviation. Our second result is that PPP disequilibria have become shorter-lived for some exchange rates, which may be due to an increase in the trade openness of the countries involved.

Key words: purchasing power parity, Markov regime-switching, testing, forecasting, exchange rates.

JEL classification: F31, C52, C53.

[*]Correspondence to: Franc Klaassen, Department of Economics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, the Netherlands; tel: +31-13-4668229; fax: +31-13-4663280; E-mail: F.J.G.M.Klaassen@kub.nl.

1 Introduction

Purchasing Power Parity (PPP) is one of the oldest and most important theories in international economics. It is commonly used as a long-run concept in relative terms, stating that in the long-run the exchange rate is proportional to the ratio of the two countries' price levels, that is, the PPP exchange rate. Long-run relative PPP is also the version of PPP we analyze in this paper. Most economists intuitively consider this hypothesis to be true. Moreover, time plots of exchange rates PPP rates support it. For instance, figures 1A, 2A and 3A plot the U.S. dollar price of one German mark, Japanese yen and U.K. pound, respectively, and the corresponding PPP rates from April 1974 to July 1997; the figures suggest a long-run comovement of exchange rates and PPP rates.[1]

Quite surprisingly, however, the existing empirical literature is not very supportive for PPP. The main contribution of this paper is that we find evidence in favor of PPP in all three of the world's most important exchange rates, that is, the U.S. dollar vis-à-vis the German mark, Japanese yen and U.K. pound. This result is not only of theoretical importance. It also has policy implications, because the long-term behavior of exchange rates relative to prices affects many economic decisions. For instance, banks need to assess the real future value of their long-term loans denominated in foreign currency. Moreover, international firms who have to decide upon foreign direct investments require reliable forecasts of the real value of the long-lasting income stream generated by the investment projects.

The reason behind the remarkable difference between our evidence in favor of PPP and the inconclusive results in the existing literature is that we use a new test. It is embedded in a model that explicitly allows for long swings in exchange rates, such as the appreciation swing of the dollar in the first half of the eighties and the subsequent depreciation swing. A long swings model seems more appropriate than the popular random walk, because it provides a better specification of the exchange rate generating process, as earlier studies have shown (see Engel and Hamilton (1990), Engel (1994) and Klaassen (1999), among others). Given the long swings model, we can test for PPP by examining whether a swing is likely to end when the PPP disequilibrium becomes large and whether the next swing governs the exchange rate back to its PPP equilibrium. More formally, we test the null of no PPP in a Markov regime-switching model in which the regime-switching probabilities depend on the PPP deviation.

Given the evidence in favor of PPP, we can go one level deeper into the concept of

---

[1] Details on the construction of the PPP rates will be given in subsection 3.1.

PPP. The argument behind PPP is that goods arbitrage equalizes prices in the same currency across countries. Because it is commonly believed that goods markets have become more integrated, making arbitrage easier, it is interesting to examine whether PPP disequilibria have become shorter-lived, which is the second and final purpose of the paper. We conclude that they have for the German mark and the U.K. pound, but not for the Japanese yen. This may indeed be explained by changes in trade openness, as we find that both European economies have become much more open, while Japan is still relatively closed.

In the literature so far, many authors have examined PPP (see Froot and Rogoff (1996) for a detailed overview). They usually concentrate on the real exchange rate and employ unit-root tests to examine the null that it follows a random walk against the alternative of stationarity, that is, PPP. Early studies, such as Meese and Rogoff (1988) and Mark (1990), use post-Bretton-Woods time series data and find no evidence of PPP. This may, of course, be due to the absence of PPP. However, it may also be caused by a lack of power of the tests. As Frankel (1986, 1990) shows, a potential reason for this is that the post-Bretton-Woods period may be too short to contain enough episodes of divergence from and reversion to PPP, because PPP disequilibria may dampen very slowly. Another cause for a lack of power may be the types of tests that are used; maybe unit-root tests are not powerful enough.

The suggestion by Frankel (1986, 1990) has resulted in two approaches to increase the power of the tests. First, Frankel (1986) and Abuaf and Jorion (1990), among others, use very long time series, often extending to a century. They indeed find evidence in favor of PPP. There is, however, some concern with these results, since the long-horizon time series blend fixed and floating rate data, and it is well-known that real exchange rates behave very differently under different exchange rate regimes (see Mussa (1986)). This is why we use only post-Bretton-Woods data.

A second way to gain power, while using only floating data, is to analyze a panel of many countries. Two notable studies in this field are Frankel and Rose (1996) and Papell (1997), which both find evidence in favor of PPP. Recently, however, O'Connell (1998) reports that the panel evidence disappears if one takes account of the strong cross-sectional dependence in real exchange rates. This argument does not apply to our results, as we analyze three exchange rates univariately.

In summary, Frankel's (1986, 1990) suggestion that PPP tests may lack power because of the use of relatively short post-Bretton-Woods time series has not resulted in conclusive evidence of PPP, despite the enormous number of studies motivated by this suggestion.

In the present paper we start from a different point of view. As mentioned above, the second potential reason for a lack of power of PPP tests so far is that they are unit-root tests, which may lack power. We investigate this claim by introducing an alternative test for PPP. Our test is based on the Hamilton (1989) regime-switching model. The basic idea of that model is that the expected change of a variable is one of two constants, depending on the regime the process is in. Persistence of such "mean regimes" leads to long swings in that variable. Because there is significant evidence of such swings in exchange rates (for instance, see Engel and Hamilton (1990), Engel (1994) and Klaassen (1999), who reject the random walk in favor of the more general long swings model), regime-switching models are a useful tool in exchange rate modeling.

A common feature of existing regime-switching exchange rate models is that they do not take account of PPP. They assume that the probability of switching to the, say, depreciation regime is constant over time, whereas PPP implies that such a switch becomes more likely when the currency is overvalued regarding its PPP value. Thus, to develop a test for PPP, we first extend the basic regime-switching model by allowing the regime-switching probabilities to depend on the PPP deviation.[2] We then derive three parameter restrictions under which the extended model implies PPP, and we test the joint validity of these restrictions. Because this test clearly supports PPP, the reason for the insignificant results from the unit-root tests in the existing literature is not the absence of PPP, but rather a lack of power.

In the next section, we introduce our regime-switching model. In section 3 we describe the data and present our empirical results. Section 4 concludes.

2 Regime-Switching Model

In this section we develop the regime-switching model that we use to answer the two questions of the paper, namely whether PPP holds and, if so, whether PPP disequilibria have become shorter-lived. We first set out the basic principles in an intuitive way. In subsections 2.1, 2.2 and 2.3 we then formally develop the model in three stages, where each stage generalizes the previous one.

The basic idea of our approach is that exchange rates exhibit two types of long swings, for instance, an appreciation and a depreciation swing. A random process governs the switches between the swings (or regimes). This regime-switching process is crucial, as the variants of the model in subsections 2.1, 2.2 and 2.3 differ with respect

---

[2]Time-varying switching probabilities are also useful when modeling switches between recessions and recoveries in the business cycle, see Durland and McCurdy (1994), Filardo (1994) and Ghysels (1994).

to this process only.

In the simplest model, see subsection 2.1, the probability of switching from one regime to the other is constant over time. Hence, the level of the exchange rate is irrelevant for the switching probabilities.

In subsection 2.2 we generalize this assumption, because it contradicts with PPP. After all, PPP implies that switches to the, say, depreciation regime are more likely when the currency is overvalued regarding PPP; the exchange rate is pulled towards its PPP equilibrium. To allow for this pull, we let the regime-switching probabilities depend on the PPP deviation. Interestingly, it appears that long-run relative PPP holds, if the pull is present and if the swing-specific appreciation and depreciation are strong enough compared to the PPP rate change (so that the exchange rate is able to return to its PPP rate after an under or overvaluation, respectively). Hence, to test for PPP, we can test whether these three conditions are fulfilled.

To answer the second question of the paper, about the duration of PPP disequilibria, we need one further generalization. It is based on the idea that PPP ensures that the long swings are around the PPP equilibrium, so that PPP disequilibria become shorter-lived if the long swings around PPP get shorter. Therefore, in subsection 2.3 we allow for a change in the duration of the swings and describe how to test whether this change is negative.

In the remaining part of this section, we formally work out the intuition just given.

2.1 Regime-Switching Model Without PPP

The regime-switching model without PPP is based on Hamilton (1989). The main difference with the basic Hamilton model is that we allow for conditional heteroskedasticity, which is present in the weekly data we use in the empirical application.

To describe the model, we need the following notation. Let S_t denote the logarithm of the spot exchange rate at time t, that is, the domestic currency price of one unit of foreign currency. We concentrate on the exchange rate change $s_t = 100(S_t - S_{t-1})$, so that s_t is the percentage depreciation of the domestic currency from time $t-1$ to t.

The regime-switching model consists of four elements. Two of them, the regime process and the mean equation, are crucial for interpreting our empirical results, as they are directly related to the exchange rate swings. The other two, the variance and distribution, will be discussed at the end of this subsection.

The regime process is based on two (unobservable) regimes. Let $r_t \in \{1, 2\}$ denote the regime at time t. Within this regime, the mean exchange rate change is μ_{r_t}, which we assume to be constant over time. Across regimes, however, the means are allowed to

differ, and we identify the first regime as the low mean regime: $\mu_1 \leq \mu_2$. This provides the basis for the swings. After all, being in the first and then in the second regime for a while leads to a period of appreciation followed by depreciation, that is, to swings in the exchange rate.

Whether swings are long or not depends on the regime staying probabilities. Let $p_{t-1}(r_t|\widetilde{r}_{t-1}) = p(r_t|I_{t-1}, \widetilde{r}_{t-1})$ denote the probability of going to regime r_t at time t conditional on the information set of the data generating process, which consists of two parts. The first part, I_{t-1}, denotes the information that is observed by the econometrician; in this subsection I_{t-1} consists of $(s_{t-1}, s_{t-2}, \dots)$. The second part, $\widetilde{r}_{t-1}$, is the regime path $(r_{t-1}, r_{t-2}, \dots)$, which is not observed by the econometrician. Note that we use the subscript $t-1$ below an operator (probability, expectation or variance) as short-hand notation for conditioning on I_{t-1}.

As in the Hamilton (1989) model, we assume in this subsection that r_t follows a first-order Markov process with constant staying probabilities, so that

$$p_{t-1}(r_t|\widetilde{r}_{t-1}) = p(r_t|r_{t-1}) = \begin{cases} p_{11} & \text{if } r_t = r_{t-1} = 1 \\ p_{22} & \text{if } r_t = r_{t-1} = 2. \end{cases} \tag{1}$$

Hence, if p_{11} and p_{22} are high, regimes are persistent and exchange rate swings are long. Note that in (1) the conditional probability that the current regime is low or high depends on the past (I_{t-1} and $\widetilde{r}_{t-1}$) only through the most recent regime r_{t-1}. This assumption represents the only difference between the current model and its generalizations in the next two subsections.

Whereas persistence in mean regimes is supposed to take account of the long swings, or "long-run autocorrelation", there may still be short-run dynamics within a mean regime. In the conditional mean specification we take account of this "short-run autocorrelation" by using one autoregressive term, as it is generally believed that the short-run autocorrelation in exchange rates is small (see West and Cho (1995)):

$$s_t = \mu_{r_t} + \theta(s_{t-1} - \mu_{r_{t-1}}) + \varepsilon_t, \tag{2}$$

where the conditional expectation of the innovation is $E_{t-1}\{\varepsilon_t|\widetilde{r}_t\} = 0$.

The regime process and conditional mean just specified are the most important elements of the model. For a complete model specification, however, we also have to define the two other elements, namely the conditional variance and distribution. This is done in the remaining part of this subsection.

When specifying the conditional variance of the error term in (2), $V_{t-1}\{\varepsilon_t|\widetilde{r}_t\}$, we take account of the conditional heteroskedasticity in the weekly data that we use in

5

the empirical application. We use the following generalized autoregressive conditional heteroskedasticity (GARCH) type model (see Bollerslev, Chou and Kroner (1992) for an overview of GARCH in standard, one-regime models):

$$V_{t-1}\{\varepsilon_t|\widetilde{r}_t\} = V_{t-1}\{\varepsilon_t\} = \omega + \alpha E_{t-1}\{\varepsilon_{t-1}^2\} + \beta V_{t-2}\{\varepsilon_{t-1}\}. \tag{3}$$

Note that we set $V_{t-1}\{\varepsilon_t|\widetilde{r}_t\}$ equal to its expectation conditional on only I_{t-1}, $V_{t-1}\{\varepsilon_t\}$. This is only for the sake of estimation simplicity.[3] We admit that it is a restriction. However, the only purpose of the variance specification is to make the PPP results, which we are mainly interested in, robust to conditional heteroskedasticity. Subsection 3.4 shows that (3) is sufficient for that.

We complete the conditional variance specification by imposing the usual GARCH restrictions $\omega > 0$ and α, $\beta \geq 0$ to ensure $V_{t-1}\{\varepsilon_t\} > 0$ for all t. We also assume that $\alpha + \beta < 1$, so that the unconditional variance is $\sigma^2 = \frac{\omega}{1-\alpha-\beta}$.

The fourth and final element of our model, the conditional error distribution, is specified by a t-distribution, which is often used to allow for extra leptokurtosis (see Bollerslev, Chou and Kroner (1992)). It has ν degrees of freedom, zero mean, and variance $V_{t-1}\{\varepsilon_t\}$:

$$\varepsilon_t \,|\, I_{t-1}, \widetilde{r}_t \quad \sim \quad t\left(\nu\,,\,0\,,\,V_{t-1}\{\varepsilon_t\}\right). \tag{4}$$

This completes the regime-switching model without PPP; it is given by (1) to (4).

2.2   Regime-Switching Model With PPP

In this subsection we extend the model of the previous subsection, so as to be able to test whether (long-run relative) PPP holds. We first examine the implications of PPP for the model and show why the model needs some extension. The required extension turns out to deal with the regime-staying probabilities in (1) only. Having described the implications of PPP, we then show that these implications also imply PPP, so that a test on their joint validity delivers a test for PPP. Finally, we give the test statistic that we will use in the empirical study.

---

[3]If we had not set $V_{t-1}\{\varepsilon_t|\widetilde{r}_t\} = V_{t-1}\{\varepsilon_t\}$, the variance would have been $V_{t-1}\{\varepsilon_t|\widetilde{r}_t\} = \omega + \alpha\varepsilon_{t-1}^2 + \beta V_{t-2}\{\varepsilon_{t-1}|\widetilde{r}_{t-1}\}$ and would have depended on the entire regime path up to time $t-1$. After all, r_{t-1} and r_{t-2} would have affected the variance through the surprise term ε_{t-1}^2, which is $\{s_{t-1} - [\mu_{r_{t-1}} + \theta(s_{t-2}-\mu_{r_{t-2}})]\}^2$ expressed in the conditioning variables, and earlier regimes would have affected the variance through the earlier surprise terms implicitly present in the lagged variance term. This would have rendered estimation intractable, since the number of possible regime paths is enormous and all regime paths have to be integrated out for estimation, as they are not observed. Specification (3) circumvents this problem by directly averaging out the regimes r_{t-1} and r_{t-2} in the source of the path-dependence, ε_{t-1}^2. The basic idea of this technique originates from Gray (1996a) and is further discussed by Klaassen (1999).

According to PPP, the deviation of the exchange rate S_t from the PPP exchange rate S_t^{ppp}, being the (logarithm of the) home price level over the foreign price level, is constant in the long-run. Therefore, if the current PPP deviation is higher (lower) than this constant, the PPP deviation is expected to fall (rise) in the long run.

In the regime-switching model, this has three implications. First, to make a fall in the PPP deviation possible, the expected change μ_1 in the low mean regime must, of course, be smaller than the expected depreciation of the PPP rate, μ_{ppp}, say. Similarly, to make a rise in the PPP deviation possible, μ_2 must exceed μ_{ppp}; this is the second implication of PPP.

The third implication concerns the regime process. In the model without PPP, the regime-staying probabilities (1) are constant over time. However, this is unrealistic if PPP holds. After all, if the current PPP deviation is, say, higher than the long-run constant, the probability of going to the low mean regime increases, so as to swing the process back into the direction of its PPP equilibrium. Hence, a large PPP deviation makes the probability of staying in the low mean regime higher, whereas staying in the high mean regime becomes less likely.

To model this dependence of the regime-staying probabilities on the PPP deviation, $S_{t-1} - S_{t-1}^{ppp}$, we use a logit specification for simplicity:

$$p_{t-1}(r_t \,|\, \widetilde{r}_{t-1}) = \begin{cases} \Lambda(\delta_1 + \delta_{ppp}(S_{t-1} - S_{t-1}^{ppp})) & \text{if } r_t = r_{t-1} = 1 \\ \Lambda(\delta_2 - \delta_{ppp}(S_{t-1} - S_{t-1}^{ppp})) & \text{if } r_t = r_{t-1} = 2, \end{cases} \tag{5}$$

where $\Lambda(.)$ is the standard logistic distribution function.[4] For parsimony, we restrict the effect of the PPP deviation to be the same (in absolute sense) for both probabilities, so that a single parameter, δ_{ppp}, captures the effect of PPP. This parameter is positive if PPP holds, and it measures the strength with which the exchange rate is pulled towards PPP equilibrium. Note that for $\delta_{ppp} = 0$ the staying probabilities are simply $\Lambda(\delta_1)$ and $\Lambda(\delta_2)$, which correspond to p_{11} and p_{22} in (1), respectively.

So far, we have concentrated on the implications of PPP for the regime-switching model: $\delta_{ppp} > 0$ is the necessary pull towards equilibrium, and $\mu_1 < \mu_{ppp} < \mu_2$ is necessary for PPP because otherwise the exchange rate will move away from PPP even if $\delta_{ppp} > 0$. To get a test for PPP, however, we need to know what these three restrictions tell us about PPP. In appendix A we show through simulations that the restrictions imply that PPP holds. Hence, one can test the null of no PPP by testing the joint null of $\mu_1 \geq \mu_{ppp}$ or $\mu_{ppp} \geq \mu_2$ or $\delta_{ppp} \leq 0$, which is the complement of the

---

[4] As opposed to the model without PPP, the information set of the econometrician, I_{t-1}, now consists of the previous exchange rate and PPP rate levels. As before, the information of the data generating process also contains the regime path.

three restrictions mentioned above. Given the existing literature, as described in the introduction, this is a new way to test for PPP. In the remaining part of this subsection, we develop the test statistic we use in subsection 3.2 of our empirical study.

We assume for simplicity that the expected PPP depreciation, μ_{ppp}, is given. This makes the null only depend on the vector $\pi = (\mu_1, \mu_2, \delta_{ppp})'$ of parameters of the regime-switching model. Since the null consists of several inequality constraints on π, we define our test statistic along the lines of Kodde and Palm (1986). That is, we use the distance from the data, represented by the maximum likelihood (ML) estimate $\widehat{\pi}$ of π, to the closest feasible point under the null (see appendix C for a description of the ML estimation procedure). More formally, our PPP test statistic is

$$\widehat{\Pi} = \min_{\pi \in H_0} (\widehat{\pi} - \pi)' \widehat{V}\{\widehat{\pi}\}^{-1} (\widehat{\pi} - \pi), \tag{6}$$

where H_0 is the set of feasible vectors π under the null, and $\widehat{V}\{\widehat{\pi}\}$ is the ML estimate for the variance of $\widehat{\pi}$.

Definition (6) shows that $\widehat{\Pi} \geq 0$ and that only points $\widehat{\pi} \notin H_0$ lead to $\widehat{\Pi} > 0$. To determine whether a realization of $\widehat{\Pi}$ is sufficiently positive to reject the null, we need the distribution of $\widehat{\Pi}$ under the null. However, we cannot use the theory in Kodde and Palm (1986) for that. After all, under the null of no PPP, the PPP deviation $S_{t-1} - S_{t-1}^{ppp}$ in the regime-staying probabilities (5) is non-stationary, making the distribution of $\widehat{\pi}$ and hence $\widehat{\Pi}$ potentially non-standard. Therefore, we simulate the null distribution of $\widehat{\Pi}$. Appendix B describes the simulation procedure that we use for our empirical study.

2.3 Duration of PPP Disequilibria

Having extended the basic regime-switching model with the allowance for PPP, we need one further extension to be able to examine the second issue of the paper, namely whether PPP disequilibria, being the difference between PPP deviations and their long run constant value, have become shorter-lived. Of course, this question is only relevant if PPP holds. Therefore, the current subsection is conditional on this. As in the previous subsection, we first extend the model to allow for a change in the duration of PPP disequilibria, and then we present the test that we use in the empirical study.

In the regime-switching model with PPP, the duration of PPP disequilibria changes if the duration of the swings around PPP changes. Since the latter depends on the intercepts in the regime-staying probabilities (5), we allow for a break in these intercepts:

$$p_{t-1}(r_t \,|\, \widetilde{r}_{t-1}) = \begin{cases} \Lambda(\delta_{10} + \delta_{ppp}(S_{t-1} - S_{t-1}^{ppp}) + \delta_{11} d_{t-1}) & \text{if } r_t = r_{t-1} = 1 \\ \Lambda(\delta_{20} - \delta_{ppp}(S_{t-1} - S_{t-1}^{ppp}) + \delta_{21} d_{t-1}) & \text{if } r_t = r_{t-1} = 2, \end{cases} \tag{7}$$

8

where d_t is one if time t is after the break date and zero otherwise.

To complete (7), we have to choose the break date. Of course, such a choice is rather ad hoc. However, from an economic point of view, the Louvre accord of February 22, 1987 is an interesting break date. After all, the Louvre accord exactly aimed at stabilizing exchange rates by introducing target zones, so as to prevent the long dollar swings of the years before. Therefore, negative values for δ_{11} and δ_{21} in (7) represent that PPP disequilibria have become shorter-lived after the Louvre accord.

To test whether δ_{11} and δ_{21} are negative, we use their ML-based t-values. These t-values have standard (normal) limit distributions, because $S_{t-1} - S_{t-1}^{ppp}$ is stationary in case of PPP. Hence, one can use standard inference. Subsection 3.3 presents the results.

3 Empirical Results

In this section we use the regime-switching model of section 2 to answer the two questions of this paper, namely whether relative PPP holds in the long-run and whether PPP disequilibria have become shorter-lived. First, we describe the data. Then, in subsection 3.2 we test for PPP and in 3.3 we examine the duration of PPP disequilibria. In subsection 3.4 we check the specification of our model. In the last subsection, we analyze whether taking account of PPP leads to better exchange rate forecasts than the simple random walk model.

3.1 Data

We use three U.S. dollar exchange rates, namely, the dollar vis-à-vis the German mark, the Japanese yen and the British pound. These exchange rates have been chosen because of their important role on foreign exchange markets and because they behave relatively independently, for instance, compared to dollar-EMS exchange rates. We take weekly instead of monthly or quarterly data, because Klaassen (1999) finds for the same series and model strategy that only weekly data yield enough observations to significantly distinguish a long swings process from a random walk, and because our central parameter δ_{ppp}, measuring the strength with which swings are pulled towards PPP, is only identified if there are swings. The data set contains 1,216 observations for the percentage dollar depreciations s_t from April 2, 1974 to July 22, 1997.

To construct the PPP exchange rates S_t^{ppp}, we follow most of the literature by using consumer price indices from the IMF International Financial Statistics.[5] They

---

[5] We use a cubic spline function to generate weekly PPP rates from the available monthly rates. The

have been obtained from Datastream, just as the exchange rates. In the remaining part of this subsection, we analyze the characteristics of the three exchange rates and PPP rates and use them to motivate our model specification.

In panel A of figures 1, 2 and 3, we show the behavior of the three actual and PPP exchange rates over the sample period (in U.S. dollars, not in logarithms). At first sight, exchange rates seem to be characterized by long swings. This impression is formally tested for the same data by Klaassen (1999), and he indeed finds that long swings are part of the exchange rate generating process. This motivates the use of a regime-switching model.

The figures also suggest that exchange rates swing around PPP and that the swings are likely to end when the deviation from the PPP rate is large. Therefore, it seems useful to let the regime-switching probabilities depend on the PPP deviations, as our model does.

Finally, we see from the plots that the swings for the two European currencies seem to be shorter in the second half of the sample. This shows that there may well have been a break in the duration of the swings. Our model allows for that.

In table 1 we report some descriptive statistics of the three exchange rate changes. There is significant first-order autocorrelation in the weekly German mark changes (we always use a significance level of 5%). Estimates for higher-order autocorrelations are not reported separately, but are combined in Box-Pierce type statistics $\widetilde{Q}_{10}$. They show that higher-order autoregressive terms in the mean equation (2) are unnecessary.

Table 1 also presents two autocorrelation tests for the squared exchange rate changes. Both tests point at conditional heteroskedasticity for all three series. This is why we have extended the basic Hamilton (1989) regime-switching model with GARCH specification (3) for the conditional error variance.

3.2 Does Relative Purchasing Power Parity Hold in the Long Run?

In this subsection we use the theory of 2.2 to answer the central question of the paper. That is, we compute the PPP statistic $\widehat{\Pi}$ in (6) to test the null of no PPP using the estimation results for the regime-switching model with PPP.[6]

The results for $\widehat{\Pi}$ [and its p-value] are: 5.46 [0.01] for Germany, 3.14 [0.05] for Japan,

---

interpolation method one chooses is practically irrelevant for the results, because PPP rates are much more stable than actual exchange rates. For convenience, we add a constant to the price index ratios such that the average PPP deviation is zero.

[6] Because the estimation results for the model of subsection 2.2 are similar to the ones for the model of 2.3 with the post-Louvre dummy (to be discussed below), we do not report the estimation results of the former model to save space.

and 4.61 [0.01] for the U.K.[7] Hence, we find evidence in favor of long-run relative PPP for all three U.S. dollar exchange rates over the post-Bretton-Woods period of floating. Given that the existing literature is not very supportive for PPP, it is remarkable that we find such conclusive evidence with our new test. This shows that the results in the literature so far are not due to the absence of PPP. Apparently, the unit-root type tests that are commonly used are not powerful enough given the relatively short post-Bretton-Woods data period.

The existence of PPP has important implications for the exchange rate swings. More specifically, exchange rate swings are more likely to end when the PPP deviation is large (see subsection 2.2). To illustrate this effect, figure 1D contains the ex ante probability of being in the high mean regime for Germany for the regime-switching models with and without PPP from 1981 to the beginning of 1986.[8] According to the model without PPP, the temporary upward moves between 1982 and 1985 are interpreted as signs of a regime-shift. However, some weeks later, it appears that there has been no such shift, and the ex ante probabilities become low again. The regime probabilities of the model with PPP are much less affected by the temporary upward moves in the early eighties. However, when the PPP deviation gets larger, their effect increases, showing that the long swing is about to end.

3.3   Have PPP Disequilibria Become Shorter-Lived?

From the previous subsection, we know that the PPP deviation is constant in the long run. In the short run, however, there are considerable periods in which the PPP deviation is different from this constant. In the current subsection we examine whether such PPP disequilibria have become shorter-lived, the second theme of the paper. We use the theory of subsection 2.3.

As argued in 2.3, we test for a change in the duration of PPP disequilibria by testing whether the swings around the PPP rate get shorter after the Louvre accord in 1987. More formally, we test whether the parameters δ_{11} and δ_{21} for the post-Louvre dummy

---

[7]Appendix B describes how we have simulated the p-values. It also argues that these p-values are conservative, that is, they are likely somewhat higher than the true p-values. However, in our case this is no problem, as the reported p-values are already low.

At first sight, it may be surprising that the p-values are so low given that the $\widehat{\Pi}$ are not very large. After all, for a test on a single one-sided restriction in a standard setting of all stationary variables, the asymptotic 5% critical value is $1.65^2 = 2.71$, where 1.65 is the 5% quantile of the standard normal distribution, and this value generally increases when non-stationary variables are involved. However, in our case the alternative hypothesis of PPP consists of three instead of a single one-sided restriction, and this has a negative effect on the critical value.

[8]The ex ante regime probability for time t is defined as the conditional probability that the process is in a particular regime at time t using only the information set of the econometrician at time $t - 1$, that is, I_{t-1} (see Gray (1996a)).

in (7) are negative. The results follow from table 3, which presents the estimates of all parameters in our model, as well as two benchmark models, namely the random walk and the regime-switching model without PPP. The table demonstrates that PPP disequilibria have become shorter-lived for the two European currencies, as three out of four coefficients for the post-Louvre dummy are significantly negative. However, we find no evidence of shorter PPP disequilibria for the yen.

The shorter duration of PPP disequilibria for the European currencies may be caused by attempts to stabilize exchange rates, such as the Louvre accord. Another reason may be the increased openness of countries. For instance, the ratio of total trade over output, which is often used as a measure for openness, has increased over our period of observation 1974-1997 from 0.42 to 0.58 for Germany and from 0.41 to 0.64 for the U.K.[9]

The shorter duration of PPP disequilibria is graphically illustrated by figures 1C and 3C, which plot the smoothed regime probabilities of being in the high mean regime for Germany and the U.K., respectively.[10] For both exchange rates we observe more, but much shorter swings after 1987, so that the exchange rates do not move far away from their PPP rates. For the U.K. the increased stability makes it even difficult to classify the observations after 1987 into one particular regime, which leads to the fairly unstable smoothed regime plot.

The second conclusion mentioned above, the lack of evidence of shorter PPP disequilibria for Japan, is in contrast with the conclusions for the two European currencies. This is, however, not surprising, because the Japanese economy is still quite closed, at least compared to Germany and the U.K., as the trade/output ratio has increased from 0.17 in 1974 to only 0.25 in 1997. This makes Japanese economic policy more independent, so that PPP disequilibria can be more persistent.

3.4 Diagnostics

The results of the two previous subsections are all based on a regime-switching model. In this subsection, we check the specification of that model in two ways, namely by

---

[9] The underlying total trade (exports plus imports) and output figures are from the OECD Main Economic Indicators in Datastream.

[10] The difference between the smoothed regime probability at time t and the ex ante probability, as defined in footnote 8, is that the former probability uses the complete data set I_T instead of only I_{t-1}, thereby smoothing the ex ante probabilities. Hence, the smoothed regime probability gives the most informative answer to the question which regime the process was likely in at time t. In appendix D we show how to compute the smoothed probabilities in a recursive manner. The algorithm is based on Gray (1996b). It links the ex ante probabilities, which are used during estimation (see appendix C), directly to the smoothed probabilities by iterating forward from the ex ante to the smoothed probabilities.

testing whether the model takes account of all autocorrelation and conditional heteroskedasticity in the data. We use the normalized residuals for that.

Table 4 presents the test results, not only for our preferred model, but also for the two benchmark models introduced before. From the first-order autocorrelations and the Box-Pierce statistics Q_{10}, we conclude that there is no remaining autocorrelation, at least for the two regime-switching models. Furthermore, the first-order autocorrelation and the aggregate autocorrelation test Q_{10}^s for the squared normalized residuals show no reason to extend the variance specifications of the three models.

3.5 Forecasting Performance

Knowing that PPP holds and that PPP disequilibria have become shorter-lived for Germany and the U.K., we now examine whether this can be exploited to predict future exchange rates better than a random walk.

We first compare the in-sample and then the out-of-sample forecasts generated by the random walk and the regime-switching model with and without PPP. We examine both point predictions and predictions of the direction of the exchange rate change by comparing the actual (logarithm of the) exchange rate level at some future time τ, S_τ, with the predicted level based on information available at time $t-1$, $\widehat{E}_{t-1}\{S_\tau\}$. For the random walk, this forecast is the previous exchange rate S_{t-1} plus an estimated drift term. For the regime-switching model, $\widehat{E}_{t-1}\{S_\tau\}$ follows from (17) in appendix E, after substitution of the estimation results of table 3. The forecasts are computed for three horizons, namely the one-week horizon, which corresponds to the data frequency, the one-quarter (13-week), and the one-year (52-week) horizon.

Starting with the in-sample forecasts, the first, often-used forecasting statistics we consider are the root mean squared error (RMSE), which is the square root of $\frac{1}{T}\sum_{t=1}^T (S_\tau - \widehat{E}_{t-1}\{S_\tau\})^2$, and the mean absolute error (MAE) $\frac{1}{T}\sum_{t=1}^T \left| S_\tau - \widehat{E}_{t-1}\{S_\tau\} \right|$. Table 5 presents their values. They show that our regime-switching model with PPP-based switching probabilities beats both the random walk and the regime-switching model without PPP in 14 out of 18 cases. The four cases where it is not the best model are all for the yen. This currency has only very few swings within our sample, so that it is not surprising that regime-switching forecasts and forecasts from a random walk are of about equal quality.

Although there is a slight preference for our regime-switching model according to the RMSE and MAE, our model clearly outperforms the other models at predicting the direction of change. In all nine cases the estimated probability of a correct prediction is higher than for the two other models. Moreover, in all cases our model predicts the

direction of change correctly in significantly more than half of the observations, while for the random walk this happens in only one case.[11] This outperformance can be attributed to two features. First, the long swings improve the forecast quality, as the regime-switching model without PPP already outperforms the random walk in eight cases. Second, the allowance for PPP in the switching probabilities leads to additional predictive power. This holds particularly at long horizons, which is in line with the fact that PPP is a long-run phenomenon.

We now turn to the out-of-sample forecasts. We reestimate the two models using only the first three quarters of the sample. Holding the parameters fixed, we then use the 304 observations in the final quarter (from November 1, 1991 to July 22, 1997) to generate the forecasts $\widehat{E}_{t-1}\{S_\tau\}$.

From table 6 we see that the superiority of our regime-switching model with PPP-based switching probabilities has vanished, at least in terms of RMSE and MAE. In only four out of eighteen cases our model outperforms both other models (in the other cases it does at least worse than the random walk). Especially for Japan our model seems to do badly. This has the same reason as given above: the swings in the yen-dollar rate are so long that there are only three switches in the in-sample period (see figure 2C). Because such switches are crucial for identifying the switching-probability parameters, the parameter estimates differ substantially from the ones based on the complete sample. Hence, more data are needed for the yen to give our model a fair chance.

Concentrating on the European currencies only, the fact that our model does not outperform the random walk may, again, be due to the rather low number of regime-switches in the in-sample period. However, it may also indicate that it is indeed difficult to beat the random walk in point prediction, as Diebold and Nason (1990) conclude from their nonparametric analysis.

Notwithstanding this result, we do find that our model outperforms the random walk at predicting the direction of change, particularly at longer horizons. The outperformance is partly due to the long swings, as the regime-switching model without PPP does already better than the random walk. Engel (1994) also reports this latter result, but he finds that the outperformance is particularly at the short-run. Our model with PPP-based switching probabilities, however, does particularly well at longer horizons,

---

[11] These conclusions about significance are robust to the autocorrelation originating from the fact that for the one-quarter and one-year horizon the forecast horizon exceeds the one week period between observations. The standard errors of the percentages are based on the Newey and West (1987) asymptotic covariance matrix. Following West and Cho (1995), we have taken Bartlett weights and have used the same data-dependent automatic lag selection rule. This rule, introduced by Newey and West (1994), has certain asymptotic optimality properties.

likely because PPP is a long-run phenomenon. The in-sample forecasting results led to the same conclusion.

4   Conclusion

In this paper we analyze the popular hypothesis of purchasing power parity (PPP), more specifically, long-run relative PPP. The main contribution of the paper is that we find evidence in favor of PPP for the three world's main U.S. dollar exchange rates over the post-Bretton-Woods period, namely the dollar vis-à-vis the German mark, Japanese yen and U.K. pound. This evidence is remarkable, because the extensive existing literature is not very supportive for PPP. The reason for this difference is that we use a new test. It is based on a regime-switching model, in which regime-switching probabilities depend on the PPP deviation.

Given the existence of PPP, we can also examine the reasons behind PPP. Our results support the view that goods arbitrage is one of the factors underlying PPP, as we find that PPP disequilibria have become shorter-lived for those countries (Germany and the U.K.) that have the largest increase in trade over the period of observation.

Our third result is that the existence of long-run PPP makes the predictions of the direction of exchange rate changes generated by our model better than those from the popular random walk model, particularly many periods ahead. The relative performance in point prediction, however, is not yet clear, because the post-Bretton-Woods data period is too short compared to the length of the swings to get sufficiently accurate in-sample estimates for the regime-switching parameters. This problem can be reduced by pooling several exchange rate series in a panel data set and then imposing some cross-sectional parameter restrictions to increase estimation accuracy. This is left for future research.

Our model can be extended and applied in various ways. Firstly, other variables such as forward rates can be included in the mean equation to improve exchange rate forecasts. Secondly, variables as the trade balance or monetary policy indicators may be informative about regime-switches, so that it may prove useful to include them besides the PPP deviation in the regime-switching probabilities. Thirdly, although we have shown that regime-switching models can provide a framework for testing long-run PPP, they may also be useful to test other long-run relationships. For instance, to test the long-run quantity theory of money, stating that the price level is proportional to the money supply in the long term. Hence, regime-switching models may offer an alternative for unit-root and cointegration tests that are commonly employed to test for long-run relations. These issues are left for future research.

Appendices

A   If $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$, then PPP holds

In subsection 2.2 we have claimed that $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$ imply PPP. In this appendix, we verify that claim. For that, we first express PPP in more formal terms.

In words, the theory of (long-run relative) PPP states that the PPP deviation is constant in the long-run. Of course, constancy is a very strict requirement. One usually means that the mean and variance of the PPP deviation are constant in the long-run and that the respective constants are independent of the current situation. We follow this interpretation. Therefore, PPP formally means that both $E_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ converge (for $\tau \to \infty$) to a limit that is independent of the conditioning information I_{t-1} and $\widetilde{r}_{t-1}$, that is, the paths of exchange rates, PPP rates and regimes up to time $t-1$.

Because we have not yet succeeded to derive a formal proof for our claim that $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$ imply PPP, we use a simulation experiment to show that it is very likely true.[12] This experiment consists of two parts. First, we demonstrate for one particular value of the initial exchange rate level S_{t-1} and the initial PPP deviation $S_{t-1} - S_{t-1}^{ppp}$, which are the only parts of I_{t-1} that we use in our simulation experiment, that under the three constraints $E_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ converge to a limit that is independent of the initial regime r_{t-1}, the only relevant part of $\widetilde{r}_{t-1}$. In the second part, we show that the two limits are also independent of the initial exchange rate and PPP deviation.

To verify the first part of our claim, we simulate both moments $E_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ for horizons one to 2,000 time periods and check our claim graphically. For that, we generate two data sets of 100,000 series of 2,000 future PPP deviations $S_\tau - S_\tau^{ppp}$. All series of both data sets start from $S_{t-1}=0$ and $S_{t-1}-S_{t-1}^{ppp}=0$, and all series within the first (second) data set are based on r_{t-1} equal to one (two). The simulated value of $E_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ ($V_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$) is defined as the mean (variance) of the 100,000 drawings of the future PPP deviation. In figures 4A and B, the two curves labeled $\delta_{ppp} > 0$ plot these simulated mean and variance, respectively, for all horizons. It is clear that $E_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ indeed

---

[12] The reported simulation results are based on the following parameter values for the exchange rate process: $\mu_1 = -0.2$, $\mu_2 = 0.2$, $\rho = 0$, $\omega = 2.5$, $\alpha = 0$, $\beta = 0$, $\nu = \infty$ and $(\delta_{10}, \delta_{11}, \delta_{20}, \delta_{21}, \delta_{ppp}) = (7, 0, 7, 0, 10)$ (the symmetry is only for the ease of interpretation). Moreover, the process for the PPP exchange rate is the random walk $s_\tau^{ppp} = 100(S_\tau^{ppp} - S_{\tau-1}^{ppp}) = \mu_{ppp} + \eta_\tau$, where $\mu_{ppp} = 0.05$ and η_τ is standard normally distributed. We have tried various other combinations, each satisfying $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$, and all yield essentially the same results.

converge to a limit that does not depend on r_{t-1} and, therefore, not on $\widetilde{r}_{t-1}$.

For comparison, figures 4A and B also contain the simulated moments in case the parameters do not satisfy the joint restrictions $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$. Since it is obvious that $E_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ do not converge if the interval of regime-specific exchange rate means does not contain the PPP trend, $\mu_{ppp} \notin (\mu_1, \mu_2)$, we only concentrate on $\delta_{ppp} \leq 0$.[13] Suppose first that $\delta_{ppp} = 0$. In that case the PPP deviation is expected to diverge, since the symmetry implied by $\mu_1 = -\mu_2$ and $\delta_{10} = \delta_{20}$ (see footnote 12) ensures that the expected exchange rate is constant in the long run, while the expected PPP rate rises. Second, the case $\delta_{ppp} < 0$ also implies a diverging PPP deviation, because moving away from the PPP rate increases the probability of staying in that situation, so that the exchange rate is expected to get stuck in one regime after a while.

In the second part of the simulation experiment, we have to demonstrate that the limits of $E_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ do not depend on the initial condition I_{t-1}, that is, on S_{t-1} and $S_{t-1} - S_{t-1}^{ppp}$, as argued before. For that, we regress 100,000 simulated values of $S_\tau - S_\tau^{ppp}$ and $(S_\tau - S_\tau^{ppp})^2$ on randomly generated S_{t-1} and $S_{t-1} - S_{t-1}^{ppp}$ and their squares for various future times τ (both initial values are generated from the uniform distribution on (-0.5,0.5)). We find that for horizons up to about 1,000 the initial condition matters, but that for longer horizons it does not.[14] Hence, the limits of $E_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ and $V_{t-1}\{S_\tau - S_\tau^{ppp}|\widetilde{r}_{t-1}\}$ indeed do not depend on I_{t-1}. Together with the conclusion from the first part of our simulation study, that both limits exist and do not depend on $\widetilde{r}_{t-1}$, this shows that $\mu_1 < \mu_{ppp} < \mu_2$ and $\delta_{ppp} > 0$ indeed imply that PPP holds.

B   P-values for PPP Tests

To decide whether the realized PPP tests $\widehat{\Pi}$ in subsection 3.2 are significant, we need the p-values. In this appendix we describe how we simulate them.

We generate 1,000 data sets, each containing one series of exchange rate levels S_t

---

[13] The reported results are based on $(\delta_{10}, \delta_{11}, \delta_{20}, \delta_{21}, \delta_{ppp}) = (4, 0, 4, 0, 0)$ and $(10, 0, 10, 0, -1)$.

[14] The White (1980) heteroskedasticity robust F-tests for no effect of S_{t-1}, S_{t-1}^2, $S_{t-1} - S_{t-1}^{ppp}$ and $(S_{t-1} - S_{t-1}^{ppp})^2$ on $S_\tau - S_\tau^{ppp}$ and $(S_\tau - S_\tau^{ppp})^2$ for horizons 100, 500, 1000, 1500 and 2000 are as follows. For $S_\tau - S_\tau^{ppp}$ as dependent variable: $2 \cdot 10^4$ [p-value is 0.00], $9 \cdot 10^1$ [0.00], 1.08 [0.36], 0.64 [0.63], and 1.38 [0.24], respectively. For $(S_\tau - S_\tau^{ppp})^2$ as dependent variable: $7 \cdot 10^3$ [0.00], $4 \cdot 10^1$ [0.00], 2.46 [0.04], 0.68 [0.61], and 1.27 [0.28], respectively.
To verify that this gradual disappearance of the effect of the initial condition is not caused by misspecification of the linear regression model, we run a nonparametric regression (see Härdle and Linton (1994)) of $S_\tau - S_\tau^{ppp}$ on S_{t-1} and $S_{t-1} - S_{t-1}^{ppp}$ separately for the horizons just mentioned. The results, which are available from the author upon request, support our claim.

and one of PPP rate levels S_t^{ppp}. Both series are generated independently ($\delta_{ppp} = 0$), so that the data satisfy the null restriction of no PPP; a detailed description of the processes underlying both series follows below. For each data set, we then estimate the regime-switching model with PPP, but without breaks in the duration of PPP deviations, thus (5) allowing for $\delta_{ppp} \neq 0$. This procedure yields 1,000 values for the PPP test statistic $\widehat{\Pi}$ in (6). The p-values for the three $\widehat{\Pi}$ that we estimate from the real data are the fraction of simulated $\widehat{\Pi}$ that exceed them.

We now discuss the 1,000 generated series of S_t and S_t^{ppp} in more detail. The length of both series is 1,217 time periods, the same as the length of the level series in the real data. Hence, our simulated p-values account for potential small-sample biases.

The process for S_t is the regime-switching process without PPP, as described in subsection 2.1. The true parameter values underlying each of the 1,000 series are the averages of the parameter estimates of the model without PPP for Germany, Japan and the U.K. (we will analyze the sensitivity of the p-values to this choice below).[15] The process starts from $S_0 = 0$.

The process for S_t^{ppp} is a normal random walk with drift. As the PPP process is independent of the exchange rate process, it is not obvious how we should choose the values of the drift parameter μ_{ppp}. After all, if μ_{ppp} equals $E\{s_t\}$, then S_t and S_t^{ppp} seem to be related through their common trend, so that one will find many large values of simulated $\widehat{\Pi}$ and thus a large p-value for the realized $\widehat{\Pi}$, so that it becomes more difficult to reject the null. On the other hand, if μ_{ppp} is outside the interval (μ_1, μ_2), then the simulated $\widehat{\Pi}$ will very often be zero, leading to a low p-value for the realized $\widehat{\Pi}$ and to an easier rejection of the null. To get an objective value for μ_{ppp}, we use the International Financial Statistics (IFS) of the IMF. We first construct the PPP rates for all countries of which the IFS contains the consumer price index. For each of these 138 PPP rates, we then estimate μ_{ppp} and the corresponding variance σ_{ppp}^2. A randomly selected pair of these two estimates is taken as the true parameter pair underlying one of the 1,000 S_t^{ppp} series. The process starts from $S_0^{ppp} = 0$.

A problem with this approach of generating S_t^{ppp} is that the OECD countries are relatively overrepresented in the IFS. Since OECD countries have quite stable PPP rates, this leads to too many μ_{ppp} close to zero. Hence, μ_{ppp} is too often close to $E\{s_t\}$, which is -0.01 for the true parameters for the S_t process.[16] As explained above, this similarity between μ_{ppp} and $E\{s_t\}$ makes the simulated p-values too high, so that it is

---

[15] As can be verified from table 3, these averages are $\mu_1 = -0.29$, $\mu_2 = 0.14$, $\theta = 0.03$, $\sigma^2 = 2.51$, $\alpha = 0.10$, $\beta = 0.88$, $\nu^{-1} = 0.19$, $p_{11} = 0.983$ and $p_{22} = 0.990$.

[16] See footnote 15, using that $E\{s_t\} = p_1\mu_1 + (1 - p_1)\mu_2$, where the unconditional regime probability $p_1 = (1 - p_{22})/(2 - p_{11} - p_{22})$, as derived by Hamilton (1989).

more difficult to reject the null.

Having described the simulation procedure, we can now present the simulated p-values that we need in the main text. The column labeled "Basic case" in table 2 shows that they are 0.00 for Germany, 0.05 for Japan, and 0.01 for the U.K. Taking account of the fact that these simulated p-values likely overestimate the true ones, we conclude that the PPP test $\widehat{\Pi}$ is significant for all three exchange rates.

A potential problem with the p-values is that they are based on one specific set of true parameters for the exchange rate process and that the p-values are likely sensitive to that choice. First, if the parameters are changed such that $E\{s_t\}$ is more similar to μ_{ppp}, then the p-values will rise, as argued above. Second, if $\mu_2-\mu_1$ is made smaller, μ_{ppp} will more often be outside (μ_1, μ_2), thereby decreasing the p-values. In the remaining part of this appendix, we demonstrate that this sensitivity indeed exists, but that it is not problematic for our conclusion of rejecting the null.

To examine the sensitivity, we compute the p-values for several combinations of nuisance parameters $E\{s_t\}$ and $\mu_2-\mu_1$, while using the same S_t^{ppp} series as before. The combinations are $E\{s_t\} = -0.1, 0, 0.1$ with $\mu_2-\mu_1$ held constant at 0.4 (to study the sensitivity regarding $E\{s_t\}$), and $\mu_2-\mu_1 = 0.2, 0.4, 0.6$ with $E\{s_t\}$ constant at 0 (to study the sensitivity with respect to $\mu_2-\mu_1$). These seem reasonable values given the estimates for the model without PPP in table 3, which imply that $(\widehat{E}\{s_t\}, \widehat{\mu}_2-\widehat{\mu}_1)$ is $(0.01, 0.42)$ for Germany, $(-0.05, 0.43)$ for Japan, and $(0.01, 0.44)$ for the U.K.

Table 2 reports the p-values corresponding to each combination of nuisance parameters. It is clear that the p-values are indeed sensitive to both $E\{s_t\}$ and $\mu_2-\mu_1$. However, the results also show that this sensitivity is not problematic for our rejection of the null of no PPP. That is, even in the worst case the largest simulated p-value (for Japan) is quite small (0.09), particularly if one takes into account that the simulated p-values overestimate the true ones, as argued above.

C   Estimation

We estimate the regime-switching model introduced in section 2 by maximum likelihood. In this appendix, we derive the likelihood function and show that it has a convenient recursive structure.

To obtain the likelihood function, we first need the density of the exchange rate change at time t conditional on observable information. Let $p_{t-1}(s_t)$ denote this density

evaluated at an exchange rate change equal to s_t.[17] It can be split up as

$$p_{t-1}(s_t) = \sum_{r_t, r_{t-1}=1,2} p_{t-1}(s_t \,|\, r_t, r_{t-1}) \cdot p_{t-1}(r_t, r_{t-1}). \tag{8}$$

We now discuss how to compute both terms on the right-hand-side.

The first term, $p_{t-1}(s_t \,|\, r_t, r_{t-1})$, denotes the density of the exchange rate change at time t evaluated at the value s_t conditional on I_{t-1} and on the current and previous regimes having values r_t and r_{t-1}, respectively. This t-density follows from formulas (2), (3) and (4). It is, however, not straightforward how to compute the conditional variance in (3), as this requires integrating out the regimes r_{t-1} and r_{t-2} in $\varepsilon_{t-1}^2 = \{s_{t-1} - [\mu_{r_{t-1}} + \theta(s_{t-2} - \mu_{r_{t-2}})]\}^2$. For that, we need $p_{t-1}(r_{t-1}, r_{t-2})$, the conditional probability that the two most recent regimes have values r_{t-1} and r_{t-2}. This probability is crucial, since all regime probabilities in the paper can be derived from it. Using similar techniques as in Gray (1996a), we now show that this probability has a first-order recursive structure, which simplifies its computation a lot.

First, we write $p_{t-1}(r_{t-1}, r_{t-2})$ as

$$
\begin{aligned}
p_{t-1}(r_{t-1}, r_{t-2}) &= p_{t-2}(r_{t-1}, r_{t-2} \,|\, S_{t-1}^{ppp}, s_{t-1}) \\
&= p_{t-2}(r_{t-1}, r_{t-2} \,|\, s_{t-1}) \cdot \frac{p_{t-2}(S_{t-1}^{ppp} \,|\, r_{t-1}, r_{t-2}, s_{t-1})}{p_{t-2}(S_{t-1}^{ppp} \,|\, s_{t-1})}.
\end{aligned} \tag{9}
$$

This equation can be simplified by assuming that the ratio on the right-hand-side is one. That is, the information contained in the two recent exchange rate regimes is irrelevant for the distribution of S_{t-1}^{ppp} once all PPP levels through $t-2$ and all exchange rate levels through $t-1$ are known. This is reasonable, since the price levels underlying S_t^{ppp} are almost fixed in the short run. Given this assumption, we have

$$
\begin{aligned}
p_{t-1}(r_{t-1}, r_{t-2}) &= \frac{p_{t-2}(s_{t-1} \,|\, r_{t-1}, r_{t-2}) \cdot p_{t-2}(r_{t-1}, r_{t-2})}{p_{t-2}(s_{t-1})} \\
&= \frac{p_{t-2}(s_{t-1} \,|\, r_{t-1}, r_{t-2}) \cdot p_{t-2}(r_{t-1} | r_{t-2}) \cdot \sum_{r_{t-3}=1,2} p_{t-2}(r_{t-2}, r_{t-3})}{p_{t-2}(s_{t-1})}.
\end{aligned} \tag{10}
$$

Hence, the variables to compute $p_{t-1}(r_{t-1}, r_{t-2})$ are its previous values $p_{t-2}(r_{t-2}, r_{t-3})$ for $r_{t-3} = 1, 2$, the previous switching probability $p_{t-2}(r_{t-1} | r_{t-2})$ and the previous densities $p_{t-2}(s_{t-1} \,|\, r_{t-1}, r_{t-2})$ and $p_{t-2}(s_{t-1})$. This makes the computation of $p_{t-1}(r_{t-1}, r_{t-2})$ a first-order recursive process.

---

[17] We use the same symbol p_{t-1} for several densities (see (1) and (8)). The specific meaning of p_{t-1} is uniquely determined by the symbols we use in its argument. This results in a concise notation, which will prove useful in the remaining part of the paper.

The second term on the right-hand-side in (8), $p_{t-1}(r_t, r_{t-1})$, is the conditional probability that the current and previous regimes have values r_t and r_{t-1}, respectively. It can be calculated from

$$p_{t-1}(r_t, r_{t-1}) = p_{t-1}(r_t \,|\, r_{t-1}) \cdot \sum_{r_{t-2}=1,2} p_{t-1}(r_{t-1}, r_{t-2}), \tag{11}$$

where $p_{t-1}(r_t \,|\, r_{t-1})$ follows from (7) and $p_{t-1}(r_{t-1}, r_{t-2})$ is given by (10).

Having discussed both terms on the right-hand-side of (8), we can now compute the density of interest, $p_{t-1}(s_t)$, being a mixture of four t-densities. This density can then be used to build the sample log-likelihood $\sum_{t=1}^{T} \log(p_{t-1}(s_t))$ with which all parameters in the regime-switching model can be estimated.

From a practical point of view, it is important to realize that the log-likelihood has a second-order recursive structure, similar to that of a standard one-regime AR(1)-GARCH(1,1) model. After all, for (11) one needs the current regime-switching probability $p_{t-1}(r_t \mid r_{t-1})$ and the first-order recursive probability $p_{t-1}(r_{t-1}, r_{t-2})$ for all eight combinations of (r_t, r_{t-1}, r_{t-2}); density (8) can then be computed from (11), the previous changes s_{t-1} and s_{t-2}, (10) and the previous variance $V_{t-2}\{\varepsilon_{t-1}\}$ in (3). This second-order recursiveness of $p_{t-1}(s_t)$ makes the calculation of the log-likelihood quite fast. To start up the recursive computation of the log-likelihood, we set the required variables equal to their expectation without conditioning on the information set.

D   Regime Inference

As stated in footnote 10, the smoothed probability that the regime was r_t at time t, $p_T(r_t)$, can be computed recursively. More generally, any ex post regime probability $p_\tau(r_t)$, for a given future time $\tau \in \{t, t+1, \ldots, T\}$, can be calculated in a recursive manner. This claim, which we prove in this appendix, is based on the following recursive process for the two-regime ex post probability $p_\tau(r_t, r_{t-1})$ starting from the ex ante probability $p_{t-1}(r_t, r_{t-1})$.

Using an assumption similar to the one below (9), we can write $p_\tau(r_t, r_{t-1})$ for the four regime combinations as

$$\begin{aligned}
p_\tau(r_t, r_{t-1}) &= p_{\tau-1}(r_t, r_{t-1} \,|\, s_\tau) \\
&= \frac{p_{\tau-1}(s_\tau | r_t, r_{t-1}) \cdot p_{\tau-1}(r_t, r_{t-1})}{\sum_{r_t, r_{t-1}=1,2} p_{\tau-1}(s_\tau | r_t, r_{t-1}) \cdot p_{\tau-1}(r_t, r_{t-1})}.
\end{aligned} \tag{12}$$

Suppose first that $\tau = t$. Then $p_\tau(r_t, r_{t-1})$ follows directly from (12), as $p_{\tau-1}(r_t, r_{t-1})$ and $p_{\tau-1}(s_\tau | r_t, r_{t-1})$ are known from the estimation process (see appendix C).

Let us suppose from now on that $\tau > t$. The computation of (12) requires two inputs. The first one is the previous ex post probability $p_{\tau-1}(r_t, r_{t-1})$, which is known from the previous recursion for all combinations of r_t and r_{t-1}. The second ingredient of (12) is the density $p_{\tau-1}(s_\tau | r_t, r_{t-1})$ for all regime outcomes. Its computation requires a number of steps. We first write it as

$$p_{\tau-1}(s_\tau | r_t, r_{t-1}) = \sum_{r_\tau, r_{\tau-1}=1,2} p_{\tau-1}(s_\tau | r_\tau, r_{\tau-1}) \cdot p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1}), \qquad (13)$$

where we use that the conditional distribution of s_τ given $r_\tau, r_{\tau-1}$ does not depend on the earlier regimes r_t and r_{t-1}. This formula itself has two ingredients. The first one is the density $p_{\tau-1}(s_\tau | r_\tau, r_{\tau-1})$ for all regime combinations, which is known from the estimation process.

The second term needed in (13) is the $(\tau - t)$-period-ahead regime-switching probability $p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1})$ for all regime combinations. Once it has been computed, it should be saved, since it will be needed in the next recursive step. Making use of the Markov structure of the regime process, it can be written in terms of $(\tau - 1 - t)$-period-ahead switching probabilities:

$$p_{\tau-1}(r_\tau, r_{\tau-1} | r_t, r_{t-1}) = \sum_{r_{\tau-1}, r_{\tau-2}=1,2} p_{\tau-1}(r_\tau, r_{\tau-1} | r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-1}(r_{\tau-1}, r_{\tau-2} | r_t, r_{t-1}). \tag{14}$$

Again, we have two ingredients. First, we need $p_{\tau-1}(r_\tau, r_{\tau-1} | r_{\tau-1}, r_{\tau-2})$ for all regime combinations. Due to the Markov property of the regime process, this switching probability does not depend on $r_{\tau-2}$. It equals

$$p_{\tau-1}(r_\tau, r_{\tau-1} | r_{\tau-1}, r_{\tau-2}) = p_{\tau-1}(r_\tau | r_{\tau-1}), \qquad (15)$$

which is known from the estimation process.

The second ingredient of (14) is $p_{\tau-1}(r_{\tau-1}, r_{\tau-2} | r_t, r_{t-1})$ for all regime combinations. Using an assumption similar to the one below formula (9), we get

$$
\begin{aligned}
p_{\tau-1}(r_{\tau-1}, r_{\tau-2} | r_t, r_{t-1}) &= p_{\tau-2}(r_{\tau-1}, r_{\tau-2} | r_t, r_{t-1}, s_{\tau-1}) \\
&= \frac{p_{\tau-2}(s_{\tau-1} | r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-2}(r_{\tau-1}, r_{\tau-2} | r_t, r_{t-1})}{\sum_{r_{\tau-1}, r_{\tau-2}=1,2} p_{\tau-2}(s_{\tau-1} | r_{\tau-1}, r_{\tau-2}) \cdot p_{\tau-2}(r_{\tau-1}, r_{\tau-2} | r_t, r_{t-1})},
\end{aligned}
\tag{16}
$$

where we use that the conditional density of $s_{\tau-1}$ is independent of the previous regimes r_t, r_{t-1} once $r_{\tau-1}, r_{\tau-2}$ are given. We have two ingredients. First, the conditional density $p_{\tau-2}(s_{\tau-1} | r_{\tau-1}, r_{\tau-2})$ for all regime combinations. It is known from the estimation process. Second, we need the $(\tau - 1 - t)$-period-ahead switching probability

$p_{\tau-2}(r_{\tau-1}, r_{\tau-2}|r_t, r_{t-1})$ for all regime combinations. This one was saved during the previous recursion, if $\tau > t+1$. If $\tau = t+1$, it equals one.

This completes the algorithm to compute (13), which is the second ingredient of (12). For each recursion one has to compute (16), use it together with (15) to compute (14) and use this to compute (13). Using this in (12) yields the ex post probability $p_\tau(r_t, r_{t-1})$ from $p_{\tau-1}(r_t, r_{t-1})$. Therefore, starting from the ex ante probability $p_{t-1}(r_t, r_{t-1})$ one can recursively compute the ex post probability $p_\tau(r_t, r_{t-1})$ and eventually the probability of interest $p_\tau(r_t)$.

E   Forecasting

Subsection 3.5 deals with forecasting exchange rate levels S_τ at time $t-1$, where $\tau \geq t$. This appendix explains how to compute these forecasts.

As usual, we first decompose the exchange rate forecast as

$$E_{t-1}\{S_\tau\} = S_{t-1} + \sum_{i=t}^{\tau} E_{t-1}\{s_i\}. \tag{17}$$

To calculate $E_{t-1}\{s_i\}$, we rewrite s_i by repeated substitution of lags of (2) for the lagged changes. Since the innovations have zero expectation, this yields

$$E_{t-1}\{s_i\} = \sum_{r_i, r_{t-1}=1,2} p_{t-1}(r_i, r_{t-1}) \cdot \left(\mu_{r_i} + \theta^{i-(t-1)}(s_{t-1} - \mu_{r_{t-1}}) \right), \tag{18}$$

where

$$p_{t-1}(r_i, r_{t-1}) = p_{t-1}(r_{t-1}) \cdot p_{t-1}(r_i \mid r_{t-1}), \tag{19}$$

where $p_{t-1}(r_{t-1})$ follows after summation of $p_{t-1}(r_{t-1}, r_{t-2})$ in (10) over r_{t-2}.

To compute the multi-period-ahead switching probability $p_{t-1}(r_i|r_{t-1})$ in (19), we first form the conditional one-period-ahead Markov transition matrices:

$$M_{t-1,j-1} = \begin{bmatrix} p_{t-1}(r_j=1 \mid r_{j-1}=1) & 1 - p_{t-1}(r_j=2 \mid r_{j-1}=2) \\ 1 - p_{t-1}(r_j=1 \mid r_{j-1}=1) & p_{t-1}(r_j=2 \mid r_{j-1}=2) \end{bmatrix}, \; j = t, \dots, i. \tag{20}$$

For $j=t$, the elements of $M_{t-1,j-1}$ follow from (7); for $j>t$, we approximate $M_{t-1,j-1}$ by $M_{t-1,t-1}$. The theory of Markov processes for multi-period-ahead switching probabilities then implies that

$$p_{t-1}(r_i \mid r_{t-1}) = (M_{t-1,t-1}^{i-(t-1)})_{r_i r_{t-1}}. \tag{21}$$

Having explained how to calculate (19), we can now compute (18). Computation of (18) for all i and substitution in (17) then gives the forecast of interest $E_{t-1}\{S_\tau\}$.

References

Abuaf, N. and P. Jorion (1990), "Purchasing Power Parity in the Long Run," *Journal of Finance*, 45, 157-174.

Bollerslev, T., R.Y. Chou and K.F. Kroner (1992), "ARCH Modeling in Finance," *Journal of Econometrics*, 52, 5-59.

Diebold, F.X. and J.A. Nason (1990), "Nonparametric Exchange Rate Prediction?," *Journal of International Economics*, 28, 315-332.

Durland, J.M. and T.H. McCurdy (1994), "Duration-Dependent Transitions in a Markov Model of U.S. GNP Growth," *Journal of Business & Economic Statistics*, 12, 279-288.

Engel, C. (1994), "Can the Markov Switching Model Forecast Exchange Rates?," *Journal of International Economics*, 36, 151-165.

Engel, C. and J.D. Hamilton (1990), "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?," *American Economic Review*, 80, 689-713.

Filardo, A.J. (1994), "Business-Cycle Phases and Their Transitional Dynamics," *Journal of Business & Economic Statistics*, 12, 299-308.

Frankel, J.A. (1986), "International Capital Mobility and Crowding-out in the U.S. Economy: Imperfect Integration of Financial Markets or Goods Markets?" in: R.W. Hafer (ed.), *How Open is the U.S. Economy?* Lexington: Lexington Books, 33-67.

Frankel, J.A. (1990), "Zen and the Art of Modern Macroeconomics: A Commentary," in: W.S. Haraf and T.D. Willett (eds.), *Monetary Policy for a Volatile Global Economy*, Washington DC: American Enterprise Institute for Public Policy Research, 117-123.

Frankel, J.A. and 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. and K. Rogoff (1996), "Perspectives on PPP and Long-Run Exchange Rates," in: G. Grossman and K. Rogoff (eds.), *Handbook of International Economics*, 3, 1647-1688.

Ghysels, E. (1994), "On the Periodic Structure of the Business Cycle," *Journal of Business & Economic Statistics*, 12, 289-298.

Gray, S.F. (1996a), "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," *Journal of Financial Economics*, 42, 27-62.

Gray, S.F. (1996b), "An Analysis of Conditional Regime-Switching Models," Working Paper, Fuqua School of Business, Duke University.

Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, 57, 357-384.

Härdle, W. and O. Linton (1994), "Applied Nonparametric Methods," in: R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics*, 4, Amsterdam: North-Holland, 2297-2341.

Klaassen, F.J.G.M (1999), "Long Swings in Exchange Rates: Are They Really in the Data?," CentER for Economic Reseach, Tilburg University, Discussion Paper No. 9908.

Kodde, D.A. and F.C. Palm (1986), "Wald Criteria for Jointly Testing Equality and Inequality Restrictions," *Econometrica*, 54, 1243-1248.

Mark, N.C. (1990), "Real and Nominal Exchange Rates in the Long Run; An Empirical Investigation," *Journal of International Economics*, 28, 115-136.

Meese, R.A. and K. Rogoff (1988), "Was it Real? The Exchange Rate Interest Differential Relation over the Modern Floating Exchange Rate Period," *Journal of Finance*, 43, 933-948.

Mussa, M. (1986), "Nominal Exchange Rate Regimes and the Behavior of the Real Exchange Rate," in: K. Brunner and A.H. Meltzer (eds.), *Real Business Cycles, Real Exchange Rates and Actual Policies*, New York: North Holland, 117-213.

Newey, W.K. and K.D. West (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703-708.

Newey, W.K. and K.D. West (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631-653.

O'Connell, P.G.J. (1998), "The Overvaluation of Purchasing Power Parity," *Journal of International Economics*, 44, 1-19.

Pagan, A.R. and G.W. Schwert (1990), "Alternative Models for Conditional Stock Volatility," *Journal of Econometrics*, 45, 267-290.

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

West, K.D. and D. Cho (1995), "The Predictive Ability of Several Models of Exchange Rate Volatility," *Journal of Econometrics*, 69, 367-391.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817-838.

Table 1: Moments of exchange rate returns and autocorrelation tests.

	GERMANY	JAPAN	U.K.
Mean	0.03	0.07	−0.03
Variance	2.14	2.11	2.13
Skewness	−0.14	0.53	−0.40
Excess Kurtosis	1.70	2.01	3.00
Autocorr. ρ_1	0.07* (0.03)	0.05 (0.04)	0.04 (0.04)
Autocorr. $\widetilde{Q}_{10}$	14.07 [0.17]	22.57* [0.01]	6.05 [0.81]
Autocorr. squares ρ_1^s	0.11* (0.03)	0.20* (0.03)	0.20 * (0.03)
Autocorr. squares Q_{10}^s	57.60* [0.00]	92.03* [0.00]	151.82 * [0.00]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.

The first-order autocorrelation, ρ_1, has been estimated as the slope coefficient in a regression of the change, s_t, on the first lagged change, s_{t-1}, and a constant. The standard errors are based on White's (1980) heteroskedasticity-consistent asymptotic covariance matrix.

$\widetilde{Q}_{10}$ denotes a modified Box-Pierce type statistic that combines the first ten autocorrelations. Following Pagan and Schwert (1990), it is defined as the sum of the first ten squared normalized autocorrelation estimates, where the normalizing factors are the heteroskedasticity-consistent standard errors of the autocorrelation estimates. $\widetilde{Q}_{10}$ is asymptotically χ_{10}^2 distributed.

The first-order autocorrelation in the squared changes, ρ_1^s, and the Box-Pierce type statistic Q_{10}^s are similarly defined, although without the correction for heteroskedasticity

Table 2: Simulated p-values for PPP tests and sensitivity to nuisance parameters.

	Basic case	Sensitivity analysis					
		Sensitivity to $E\{s_t\}$			Sensitivity to $\mu_2-\mu_1$		
Unconditional mean $E\{s_t\}$	-0.01	0.1	0	-0.1	0	0	0
Wedge regime means $\mu_2-\mu_1$	0.43	0.4	0.4	0.4	0.2	0.4	0.6
Critical value of PPP test $\widehat{\Pi}$	3.22	2.23	3.06	3.39	2.45	3.06	3.81
P-value Germany ($\widehat{\Pi}=5.46$)	0.00	0.00	0.00	0.01	0.00	0.00	0.01
Japan ($\widehat{\Pi}=3.14$)	0.05	0.01	0.05	0.06	0.02	0.05	0.09
U.K. ($\widehat{\Pi}=4.61$)	0.01	0.01	0.01	0.03	0.00	0.01	0.04

The column labeled "Basic case" contains the p-values that are used in the main text. These are computed from exchange rate and PPP rate processes simulated from parameter values that are equal to the average estimates of the model without PPP (see footnote 15).

The sensitivity analysis is based on different combinations of $E\{s_t\}$ and $\mu_2-\mu_1$. To transform each $(E\{s_t\}, \mu_2-\mu_1)$ into the structural parameters μ_1 and μ_2, we assume for simplicity that the unconditional regime probabilities are both 0.5, so that $\mu_1 = E\{s_t\} - 1/2(\mu_2-\mu_1)$ and $\mu_2 = E\{s_t\} + 1/2(\mu_2-\mu_1)$. This is obtained by taking $p_{11}=p_{22}$, which we set at 0.987, the average of the values in footnote 15. The other parameters are kept at the average parameter estimates of the model without PPP (see footnote 15).

Further details about the simulation procedure are in Appendix B.

Table 3: Estimation results.

		GERMANY			JAPAN			U.K.		
		RW	noPPP	PPP	RW	noPPP	PPP	RW	noPPP	PPP
Mean of regime	μ_1	0.03 (0.04)	-0.27* (0.09)	-0.29* (0.07)	0.01 (0.03)	-0.30 (0.15)	-0.36* (0.09)	0.01 (0.03)	-0.30* (0.09)	-0.31* (0.10)
	μ_2		0.15* (0.07)	0.20* (0.05)		0.13 (0.07)	0.09* (0.03)		0.14* (0.06)	0.16* (0.05)
Autocorr.	θ		0.07* (0.03)	0.06* (0.03)		0.04 (0.03)	0.05 (0.03)		-0.01 (0.03)	-0.01 (0.03)
Regime stay prob	p_{11}		0.992 (0.010)			0.976 (0.028)			0.981 (0.021)	
	p_{22}		0.996 (0.007)			0.983 (0.019)			0.992 (0.013)	
Logit intercept	δ_{10}			10.98* (3.12)			8.08* (2.56)			5.85* (1.34)
post-Louvre	δ_{11}			-7.77* (3.09)			0			-4.44* (1.65)
	δ_{20}			10.28* (2.82)			9.10* (2.96)			5.63* (1.19)
	δ_{21}			-4.93* (2.18)			0			-1.31 (1.34)
PPP deviation	δ_{ppp}			17.18 (7.29)			13.53 (7.63)			8.13 (4.23)
Uncond. variance	σ^2	2.89 (1.08)	3.11 (1.41)	3.07 (1.35)	1.82 (0.87)	1.62 (0.84)	1.80 (0.90)	2.86 (1.11)	2.81 (1.11)	2.73 (1.06)
ARCH	α	0.13* (0.03)	0.14* (0.03)	0.15* (0.03)	0.07* (0.02)	0.07* (0.02)	0.07* (0.02)	0.11* (0.02)	0.10* (0.02)	0.10* (0.02)
GARCH	β	0.84* (0.04)	0.83* (0.04)	0.82* (0.04)	0.92* (0.02)	0.92* (0.02)	0.92* (0.02)	0.88* (0.02)	0.89* (0.02)	0.89* (0.02)
T-dist.	ν^{-1}	0.12* (0.03)	0.14* (0.03)	0.14* (0.03)	0.20* (0.02)	0.21* (0.02)	0.21* (0.02)	0.20* (0.02)	0.22* (0.03)	0.21* (0.03)
Log-likelihood minus RW		-2126 0	-2116 9.34	-2110 15.82	-2053 0	-2044 8.91	-2043 10.41	-2068 0	-2062 6.34	-2057 11.28

Standard errors in parentheses; * is significant at 5% level.

"RW" denotes the random walk, "noPPP" ("PPP") the regime-switching model without (with) allowance for PPP (see (1) and (7), respectively.)

Because of our evidence in favor of PPP, the PPP deviation $S_{t-1} - S_{t-1}^{ppp}$ in (7) is stationary. Therefore, the t-values for all parameters except δ_{ppp} have the standard (normal) asymptotic distribution, so that one can use standard inference. For δ_{ppp} the t-value may well have a non-standard limit distribution, so that we do not know for sure whether the estimates in the table are significant.

The two zeros in table 3 for Japan indicate that we have to impose $\delta_{21} = \delta_{22} = 0$ to achieve convergence. This restriction is realistic, as figure 2A shows no signs of a structural break in the yen-dollar swings after the Louvre accord.

We present the inverse of the degrees of freedom of the t-distribution, because testing for conditional normality then boils down to simply testing whether ν^{-1} differs significantly from zero.

Table 4: Diagnostic statistics for normalized residuals and their squares.

	GERMANY			JAPAN			U.K.		
	RW	noPPP	PPP	RW	noPPP	PPP	RW	noPPP	PPP
Autocorr. ρ_1	0.10*	0.01	0.01	0.08*	0.01	0.01	0.06*	0.04	0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Autocorr. Q_{10}	24.40*	6.47	6.53	34.11*	17.37	18.86*	16.32	6.37	5.87
	[0.01]	[0.78]	[0.77]	[0.00]	[0.07]	[0.04]	[0.09]	[0.78]	[0.83]
Autocorr. ρ_1^s	-0.05	-0.05	-0.06	0.06*	0.06*	0.06*	0.03	0.04	0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Autocorr. Q_{10}^s	16.32	15.87	17.75	11.13	11.16	10.99	9.31	9.91	10.21
	[0.09]	[0.10]	[0.06]	[0.35]	[0.35]	[0.36]	[0.50]	[0.45]	[0.42]

Standard errors in parentheses and p-values in square brackets; * is significant at 5% level.

"RW" denotes the random walk, "noPPP" ("PPP") the regime-switching model without (with) allowance for PPP (see (1) and (7), respectively.)

The residual is the exchange rate change minus the estimate of its conditional expectation $E_{t-1}\{s_t\}$. The regime probability to integrate out the unobserved regimes in this expectation can be found in appendix C. The residual is normalized by its variance, $V_{t-1}\{s_t\}$. Note that it is not equal to the error variance $V_{t-1}\{\varepsilon_t\}$, since the possibility of regime-switches is an additional source of variance of the residuals besides the one represented by the error term.

All autocorrelation statistics have been defined below table 1, although the standard error of ρ_1 and the value of Q_{10} are no longer corrected for heteroskedasticity.

Table 5: In-sample forecasting statistics.

	GERMANY			JAPAN			U.K.		
	RW	noPPP	PPP	RW	noPPP	PPP	RW	noPPP	PPP
Panel A: One-week horizon									
RMSE	1.464	1.458	1.448	1.454	1.449	1.452	1.459	1.455	1.445
MAE	1.095	1.085	1.080	1.041	1.033	1.033	1.043	1.038	1.034
Correct direction	0.527* (0.014)	0.562* (0.014)	0.562* (0.014)	0.484 (0.014)	0.548* (0.014)	0.552* (0.014)	0.507 (0.014)	0.560* (0.014)	0.561* (0.014)
Panel B: One-quarter horizon									
RMSE	5.941	5.959	5.522	6.305	6.368	6.369	5.974	5.944	5.482
MAE	4.814	4.757	4.347	4.956	4.916	4.914	4.585	4.485	4.217
Correct direction	0.530 (0.045)	0.576* (0.041)	0.687* (0.036)	0.539 (0.047)	0.586* (0.038)	0.591* (0.045)	0.492 (0.046)	0.579* (0.039)	0.647* (0.038)
Panel C: One-year horizon									
RMSE	12.945	13.487	11.035	14.059	14.751	14.280	12.891	12.911	9.724
MAE	10.585	10.338	8.411	11.042	11.581	11.210	10.722	10.280	7.668
Correct direction	0.534 (0.065)	0.597* (0.056)	0.736* (0.049)	0.609* (0.063)	0.535 (0.049)	0.648* (0.057)	0.480 (0.065)	0.589 (0.054)	0.767* (0.046)

Standard errors in parentheses; * is significantly greater than 0.5 at 5% level.
"RW" denotes the random walk, "noPPP" ("PPP") the regime-switching model without (with) allowance for PPP (see (1) and (7), respectively.)
"Correct direction" denotes the fraction of forecasts that yield the correct direction of change of the exchange rate level. For the one-quarter and one-year horizon the standard errors have been corrected for autocorrelation as explained in footnote 11.

Table 6: Out-of-sample forecasting statistics.

	GERMANY			JAPAN			U.K.		
	RW	noPPP	PPP	RW	noPPP	PPP	RW	noPPP	PPP
Panel A: One-week horizon									
RMSE	1.523	1.526	1.526	1.511	1.515	1.544	1.465	1.473	1.463
MAE	1.133	1.136	1.139	1.097	1.099	1.098	1.000	1.006	1.012
Correct direction	0.512 (0.029)	0.531 (0.029)	0.502 (0.029)	0.454 (0.029)	0.484 (0.029)	0.539 (0.029)	0.459 (0.029)	0.502 (0.029)	0.502 (0.029)
Panel B: One-quarter horizon									
RMSE	5.612	5.680	5.810	6.490	6.562	7.643	5.638	5.759	5.506
MAE	4.589	4.663	4.575	5.106	5.026	5.935	3.671	3.784	3.928
Correct direction	0.438 (0.075)	0.486 (0.076)	0.599 (0.072)	0.503 (0.081)	0.545 (0.071)	0.575 (0.075)	0.490 (0.074)	0.483 (0.048)	0.594 (0.071)
Panel C: One-year horizon									
RMSE	10.151	11.033	13.301	12.765	12.803	20.394	9.470	9.515	8.253
MAE	8.807	9.489	11.528	11.059	10.787	17.234	6.995	7.297	7.339
Correct direction	0.455 (0.101)	0.498 (0.095)	0.573 (0.095)	0.605 (0.106)	0.628 (0.080)	0.553 (0.101)	0.522 (0.095)	0.522 (0.095)	0.684* (0.084)

Standard errors in parentheses and p-values in square brackets; * is significantly greater than 0.5 at 5% level.

"RW" denotes the random walk, "noPPP" ("PPP") the regime-switching model without (with) allowance for PPP (see (1) and (7), respectively.)

The whole series except for the last quarter has been used for estimation, while the last quarter (304 weeks from November 1, 1991 to July 22, 1997) has been used for forecasting. This means that for the one-quarter (year) horizon there are 292 (253) comparisons between the actual and predicted values.
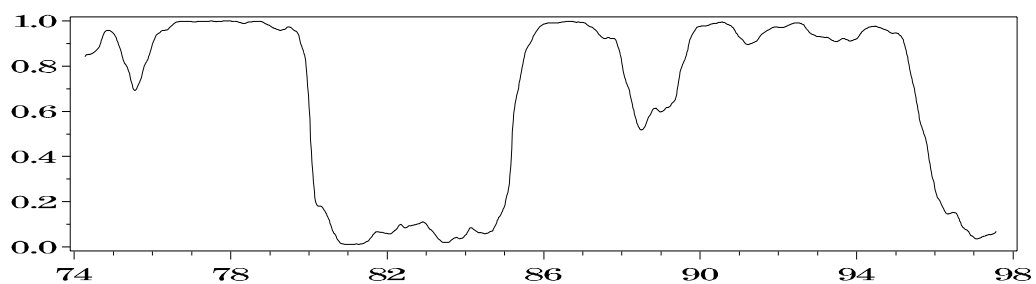
"Correct direction" denotes the fraction of forecasts that yield the correct direction of change of the exchange rate level. For the one-quarter and one-year horizon the standard errors have been corrected for autocorrelation as explained in footnote 11.
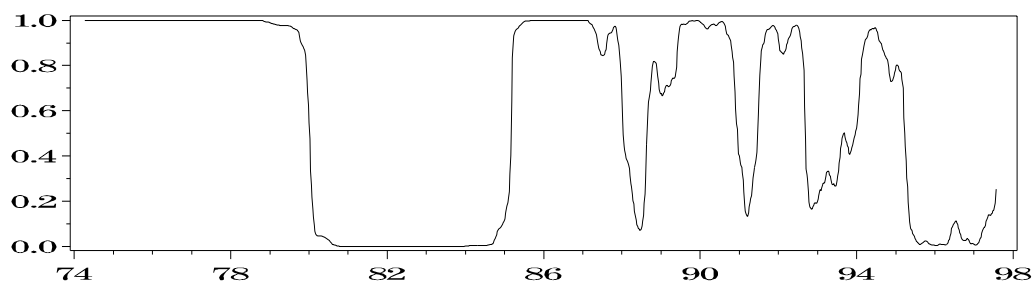
A: Exchange rate levels (in dollars) and PPP (smooth curve)
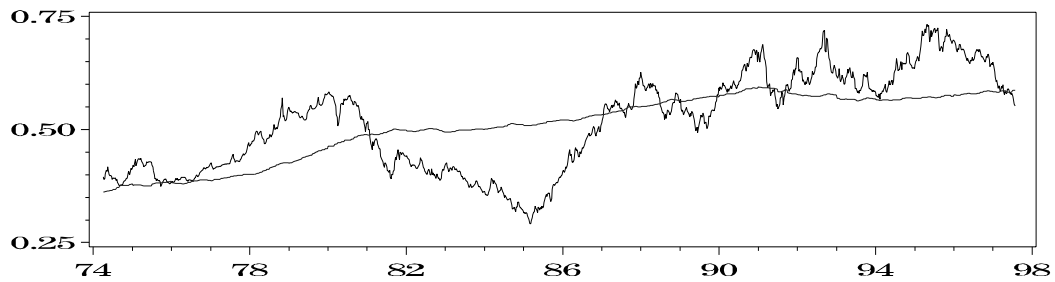
B: Smoothed regime probabilities without PPP

C: Smoothed regime probabilities with PPP

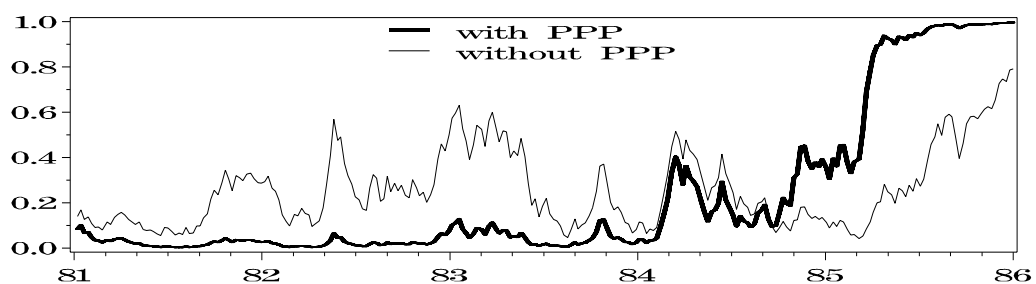D: Effect of PPP on ex ante regime probabilities from 1981 to 1986

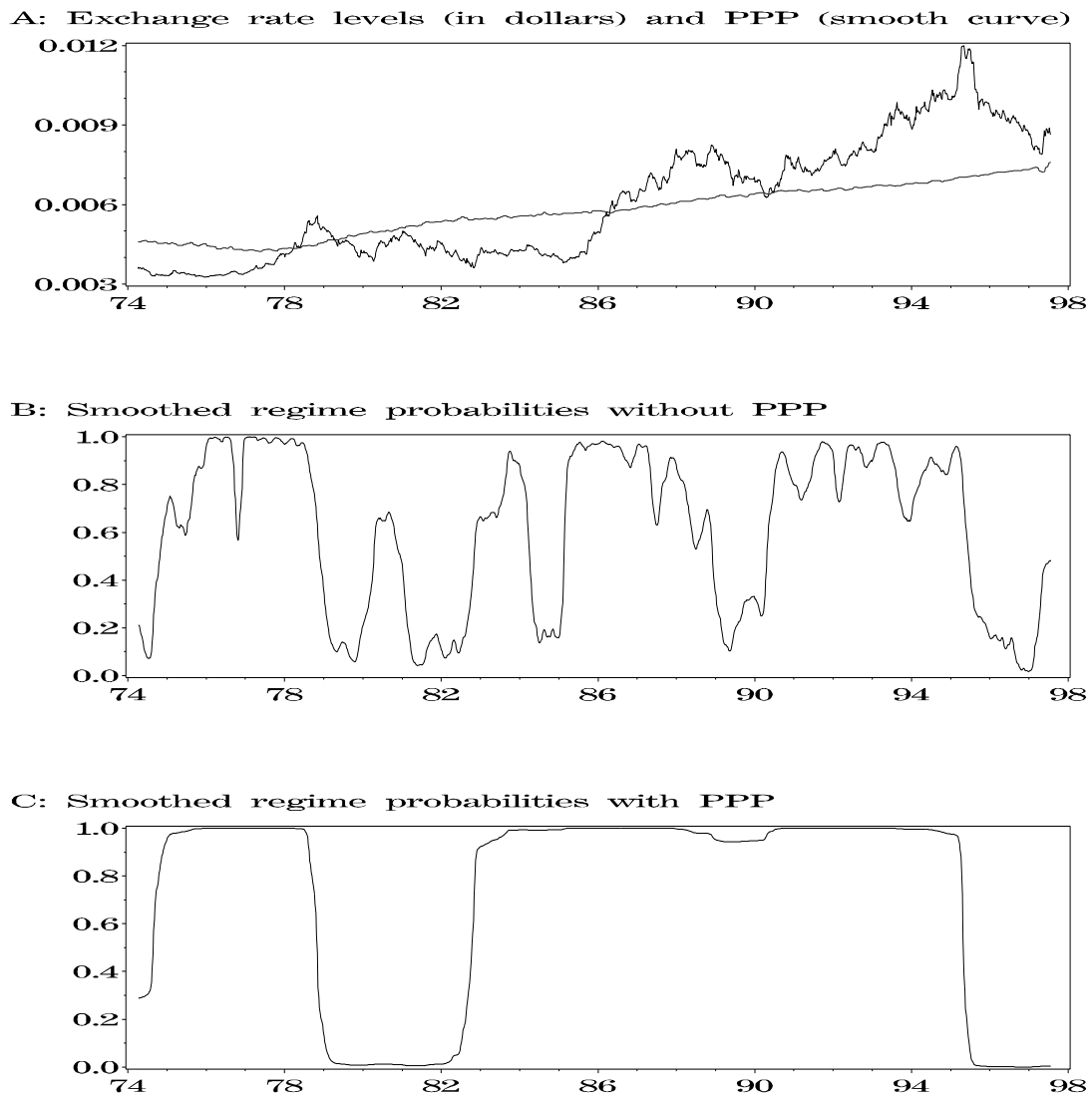Figure 1: German mark over the sample period April 1974 to July 1997.

A: Exchange rate levels (in dollars) and PPP (smooth curve)

B: Smoothed regime probabilities without PPP

C: Smoothed regime probabilities with PPP

Figure 2: Japanese yen over the sample period April 1974 to July 1997.

33

A: Exchange rate levels (in dollars) and PPP (smooth curve)

B: Smoothed regime probabilities without PPP

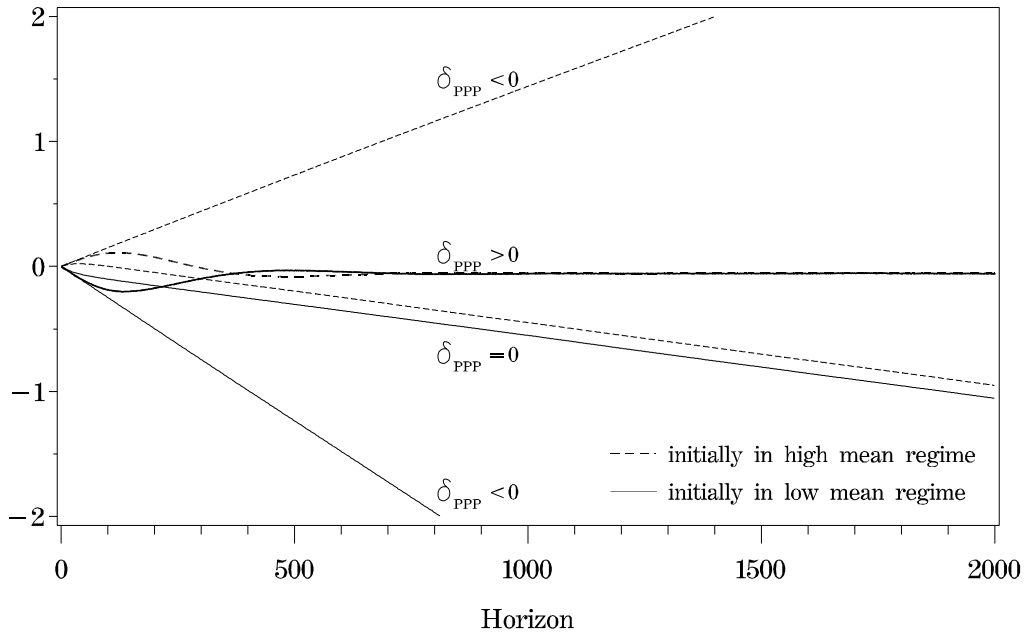**C: Smoothed regime probabilities with PPP**
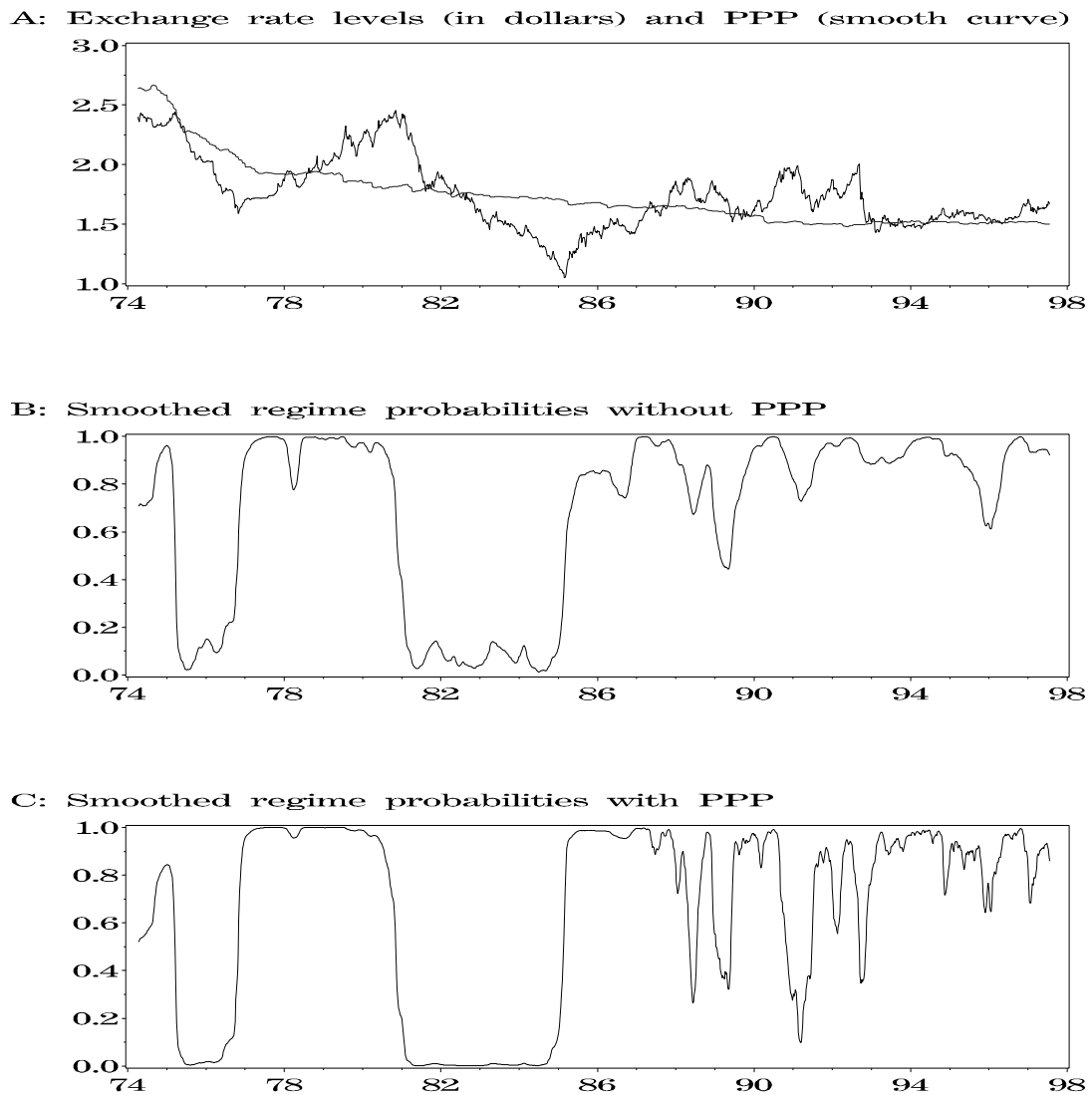
Figure 3: U.K. pound over the sample period April 1974 to July 1997.

A: Mean of future PPP deviation
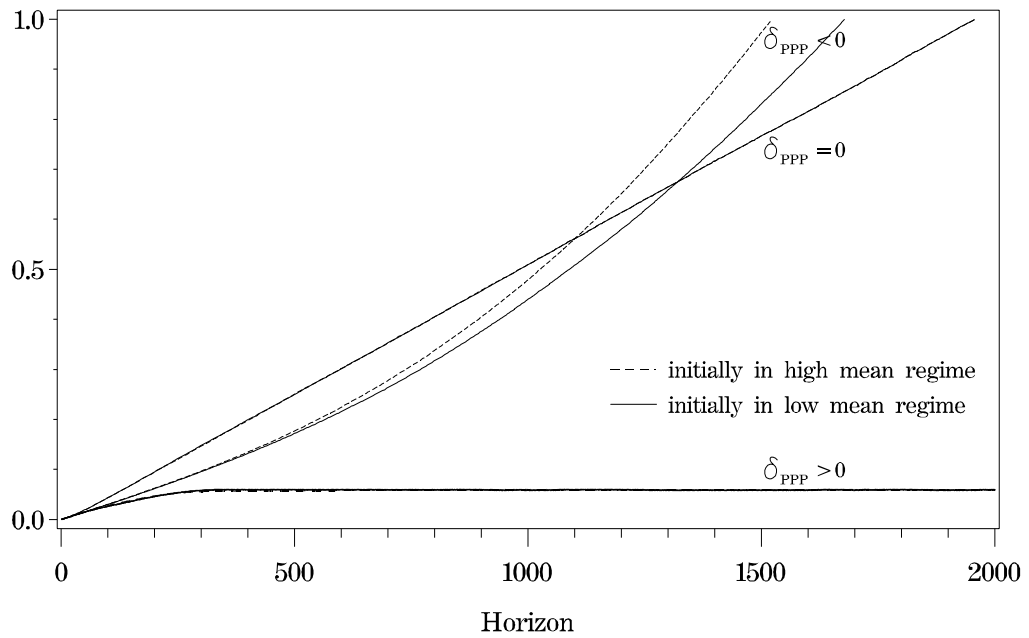

B: Variance of future PPP deviation

Figure 4: Behavior of future PPP deviations for different δ_{ppp} (measuring the strength with which the exchange rate is pulled towards PPP).