The ESTAR Model:

Nonlinearities, Price Indices, Crises and Forecasts

Master Thesis

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Department of Economics

Author:

Maurits van Kempen

343022

343022mk@student.eur.nl

Supervisor:

Dr. L. Pozzi

Abstract

Due to arbitrage, real exchange rates might revert to their equilibrium value in a nonlinear fashion. Modelling such behavior could enhance our understanding and forecasting of exchange rate movements. In this report, the Exponential Smooth Transition Autoregressive model is used on monthly data on four currency pairs using Nonlinear Least Squares in an attempt to model nonlinear exchange rate behavior. The report also focuses on the role of trade price indices, the influence of the recent financial crisis and the forecast performance of the model. Limited evidence of nonlinearity is found when comparing the test statistics with their simulated counterparts. Furthermore, no evidence of the impact of the crisis is found and using the trade price index only adds value over the traditional price index when forecasting the exchange rate.

1 Introduction

The aim of this report is to analyse a model concerning exchange rate behavior, called the Exponential Smooth Transition Autoregressive (ESTAR) model, with the help of several insights and methods from the theoretical and empirical literature. Doing so enhances our understanding of real exchange rate movements and could help predict such movements. The model is estimated for four currency pairs with monthly data over a time period of 33 years. Moreover, the effect of the recent financial crisis is analysed and the forecasting performance of the model is evaluated. This section briefly explains the subject and contents of this report.

1.1 Introduction to ESTAR

The ESTAR model is based on the idea that in the long run, some form of the concept of purchasing power parity holds. This concept, abbreviated to PPP, tells us that the price levels of two countries should equalize when the exchange rate between those countries is taken into account. If price levels are not equal, international goods traders will act to make a profit, forcing the real exchange rate to move so as to make PPP hold. This occurs because the traders can buy goods in the country with the lower price and sell them in the country with the higher price. Eventually, this process, called arbitrage, should even out supply and demand so that the real exchange rate adjusts back towards PPP.

The ESTAR model embraces this entire process of equalization, but also adds something new; it assumes that the process of the real exchange rate returning to PPP takes place *nonlinearly*. The nonlinearity in this arbitrage process comes about because bigger price differences between countries allow for bigger profits and hence attract more and faster arbitrage. Larger deviations of the real exchange rate from PPP are thus expected to be eliminated quicker than small deviations.

The idea described above can be further specified and captured in an ESTAR model equation, which can then be tested empirically. This report does exactly that, but adds to the existing literature by making use of some insights and applications from different parts of the exchange rate theory and literature. A first example is the use of the Trade Price Index (TPI) alongside the Consumer Price Index (CPI). The TPI focuses on the prices of tradable goods and services, in contrast to CPI, which also includes the prices of nontradables. By using TPI, one should avoid the case in which PPP does not hold because of differences in the price of nontradables.

Secondly, this report uses the concept of a time-varying PPP-implied equilibrium exchange rate instead of a fixed PPP equilibrium value. The concept of a fixed value is based on the strict absolute version of PPP for which empirical support generally lacks. The time-varying PPP-implied equilibrium however, is based on a weaker form of absolute PPP and on relative PPP. These two versions of PPP are far less demanding and receive more empirical support. This could make the time-varying equilibrium a more acceptable PPP benchmark.

A third example are the dummy variable analysis for the financial downturn in the late 00's and the forecasting attempt. By using a dummy variable in the ESTAR specification, this report investigates whether the latest financial crisis has had an impact on the nonlinear behavior of the real exchange rate. The rationale for such an effect would be that traders are less willing to take risks when financial turmoil is present on the world markets. This reduced arbitrage would then allow for a larger price gap between countries to persist. The forecasting attempt concerns a recursive regression of the ESTAR model to forecast the real exchange rate in the years 2000 to 2013. These ESTAR forecasts are then compared to a benchmark random walk model that embodies the uninformed forecaster.

1.2 The contents and findings

Before the aforementioned applications and estimations take place, the relevant literature on PPP and exchange rate models is discussed in section 2. In section 3 the specification of the ESTAR model, the construction of the time varying PPP-implied equilibrium and the data and used methods are covered. The estimations of the model are performed using Nonlinear Least Squares on monthly data from January 1980 to December 2013 for the U.S. dollar (the benchmark currency), the Canadian dollar, the Japanese yen, the Dutch guilder and euro and the Singapore dollar. Section 4 covers the results of these estimations. Finally, section 5 concludes. Some relevant figures and tables can be found in the Appendix at the end of this report.

In this report, weak support for ESTAR nonlinear behavior of real exchange rates is found. The estimated coefficients are of a plausible size and have the correct sign. These results are somewhat less supportive of the ESTAR model than most of the other empirical works on this topic. I find no evidence of any impact of the recent financial crisis on the nonlinear behaviour of the exchange rates. The forecasting qualities of the ESTAR models surpass that of the random walk models, and the TPI-based models forecast more accurately than the CPI-based models. The latter result is in line with another empirical work in similar vein.

2 Literature review

Generally speaking, one could classify this empirical investigation as belonging to the broad strand of research concerning purchasing power parity. However, this report covers a more specific subtopic of PPP concerning the nonlinear behavior of the real exchange rate. Even at the level of this subtopic, it is still possible to distinguish between the various different types of models that were developed for that nonlinear relationship. In this section, these layers of literature are covered separately. It begins with an introduction to purchasing power parity and the associated history of research techniques and results. Afterwards, some major reasons and theories for deviations from PPP are discussed. This section concludes with a discussion of the literature on exchange rate nonlinearities and the associated models.

2.1 The basics of PPP

Both the history and the massive bulk of literature on purchasing power parity can be called impressive. The concept of PPP was first studied by a group of scholars in sixteenth century Imperial Spain, called the Salamanca school (Einzig, 1970). The idea behind the concept stems from the law of one price which states that any good that is sold in two different currency regions must have the same price in both regions when the exchange rate is taken into account. Violations of this economic law would imply that profitable arbitrage is possible. PPP extends this no-arbitrage condition by combining all prices for all goods of a region. Thus, without any obstructions, the aggregate price level of any two currency regions is expected to be equalized when the exchange rate is taken into account:

$$P = S * P^* \tag{1}$$

where P is the domestic price level, P^* is the foreign price level and S is the spot exchange rate. This form of PPP is known as the absolute version. The relative version can be obtained by rearranging equation (1) and taking the first difference:

$$\frac{S_1 - S_0}{S_0} = \frac{P_1 - P_0}{P_0} - \frac{P_1^* - P_0^*}{P_0^*} \tag{2}$$

where the subscripts denote time so that $\frac{S_1 - S_0}{S_0}$ is the percentage (decimal) change in the spot exchange rate which equals the difference between the rates of inflation in both regions. As the topic of this paper is mostly related with the absolute version, the literature on the relative version is not covered here.

PPP lends itself for determining whether a currency is overvalued or undervalued relative to some other currency. Perhaps the simplest, yet most insightful way of doing this is The Economist's Big Mac Index, which compares the prices of McDonald's Big Mac burger across the world (The Economist, 2014). This 'easily digestible' index has served as a topic for several publications, including that of Pakko and Pollard (2003). In contrast to this modern lighthearted approach, the application of PPP was discussed in very serious matters at the end of World War I, when countries searched for a new way to value their currencies (Rogoff, 1996). Using PPP to find this value was advocated by the Swedish economist Gustav Cassel (1921). These examples show how PPP plays a major role when thinking about exchange rates and relative prices. It is therefore useful to take a closer look at the empirical research dedicated to this concept.

2.2 The empirical history of PPP

Froot and Rogoff (1984) distinguish three phases of the history on PPP research. The first phase is the period in which unit roots and random walks were unknown or ignored. Tests for purchasing power parity were executed using naïve bivariate Ordinary Least Squares (OLS) regression with the exchange rate on the left hand side and the relative price differential on the right hand side, as shown in equation (3):

$$s_t = \alpha + \beta(p_t - p_t^*) + \epsilon_t \tag{3}$$

where s_t is the natural logarithm of the nominal exchange rate, p_t and p_t^* are the natural logarithms of the home and abroad price level respectively and ϵ_t is the error term. Furthermore, α and β are the OLS coefficients. The usual null hypothesis at the time was that $\beta = 1$ so that PPP holds. Frenkel (1978) used such an approach on data from the 1920s and found significant evidence for the PPP hypothesis. However, as non-stationarity was not taken into account, it is troublesome to say anything meaningful about these early studies.

The second phase starts with the realization of this flaw and is characterized by the null hypothesis that deviations from PPP are permanent. To test this hypothesis, we need a variable that expresses the deviation from PPP. The real exchange rate does exactly this. If we impose that $\beta = 1$ (and ignore α and ϵ_t for now) in equation (3), we can rewrite it to get the expression for the deviation from PPP, as shown in equation (4):

$$q_t = s_t - p_t + p_t^* \tag{4}$$

Here, q_t is the real exchange rate, which equals the deviation from PPP: it measures the gap between the nominal exchange rate and the price differential. Now, to say that deviations from PPP are permanent is to say that q_t, the real exchange rate, contains a unit root (is nonstationary). In that case, the movement of the real exchange rate is random; it has no tendency to return to a certain mean. This extreme view is now the null hypothesis, while the alternative is that PPP holds in the long run. The latter would imply that the real exchange rate is stationary. Thus, in this phase, it all comes down to testing for stationarity. To do this, more modern time series econometrics, such as unit root tests, are employed in this phase. Using those methods, Hakkio (1984) and Huizinga (1987) are able to reject the null hypothesis of the random walk. On the contrary, Adler and Lehman (1983) and Edison (1987) find no evidence for long run purchasing power parity. Slightly more positive are the findings of Frankel (1986), who is able to reject nonstationarity with a 116 year sample. Roll (1979), on the other hand, finds that the real exchange rate is characterized by a random walk. He notes that this finding can comply with the theory if one assumes that markets are efficient. This market efficiency would make it impossible to predict the future of the exchange rate. Froot and Rogoff (1984) disagree with Roll, arguing that there is no 'market' for the real exchange rate and therefore no reason for it to behave like an asset. The authors also warn us that one should be aware that the econometric tests used in this second phase are plagued with low power. This makes it difficult to distinguish between a random walk and stationarity. In general then, the literature of the second phase indicates that, for floating currencies at least, it is difficult to reject the existence of a unit root.

In the third phase, cointegration approaches, such as the work of Engle and Granger (1991) gained popularity among PPP-researchers. The idea behind the cointegration relationship is that, although exchange rates and price indices could be random walks by themselves, a linear combination of the two series could be stationary. In the Engle-Granger procedure, this is tested by regressing the exchange rate on the relative price difference, as shown in equation (5):

$$s_t = \alpha + \beta_1 p_t - \beta_2 p_t^* + \epsilon_t \tag{5}$$

Note that both the home and foreign price level are now allowed their own coefficient: β_1 and β_2. The estimation of equation (5) produces a residual term e_t (the post-estimation counterpart of ϵ_t), which can then be tested for a unit root. If it does not contain a unit root, it is said that the exchange rate and the relative price differences are cointegrated. Froot and Rogoff (1984), however, point out that the procedure described above is inherently inefficient for PPP testing in the form of equation (5). The issue lies in the fact that equation (5) has three variables between which, in theory, *two* cointegrating relationship can exist. The Engle Granger procedure, however, only allows detecting *one* single cointegrating relationship. This is problematic since the estimates of that single resulting cointegrating relationship will depend on the choice of the left hand side variable of equation (5), while such a choice should not matter. One solution for this issue is using equation (3) instead of equation (5) for estimation, as equation (3) only has two variables. Another solution is using equation (5) with the Johansen (1991) cointegration procedure. This procedure uses a Maximum-Likelihood estimator which is not influenced by the choice of left hand side variable. This allows detecting more than one cointegrating relationship when there are three variables. This solution is preferred as it is less restrictive: using equation (5) allows each price index its own coefficient.

The mentioned works on cointegration stem from a period in which PPP studies started focusing on the long run properties of the PPP relationship. McMahon and Taylor (1988) find some evidence of long run PPP between the dollar and the pound for the start of the twentieth century. Kim (1990) also finds that PPP seems to hold generally. Giovannetti (1987) however, finds no cointegration for the dollar versus 7 other currencies. The same result goes for Taylor (1988), Corbae and Ouliaris (1988), Hakkio and Rush (1989) and Mark (1990). Still others, such as Enders (1988) and Fisher (1991) hold the middle ground with mixed evidence. So far then, the literature does not provide us with a clear answer concerning the validity of purchasing power parity. To form a bridge to the more recent literature and to some key factors driving this report, it is useful to shortly survey some theorized causes of deviations from long run PPP.

2.3 Deviations from PPP

As noted before, purchasing power parity can be seen as a reformulated and aggregated version of the law of one price. Deviations from PPP could therefore just be the aggregate of deviations from the law of one price. In the early literature on the law of one price, results were mixed. Genberg (1975) showed that the no-arbitrage concept tends to hold for heavily traded commodities, such as cocoa and rubber. However, in an influential study, Isard (1977) extended the range of products to include items such as industrial chemicals and paper and found persistent deviations from the law of one price.

2.3.1 Price indices

Most studies on PPP make use of the standard Consumer Price Index (CPI) or the Wholesale Price Index (WPI). For the calculation of the CPI, the goods and prices that are relevant for consumers are used, while WPI focuses more on intermediate goods (semi-finished-products) and prices. The calculations for any price index are based on a certain basket of goods. Even for the same 'type' of index (CPI or WPI or others) these baskets do not necessarily contain identical goods in every country and neither is the weight given to each individual good in the basket the same. If every price index is calculated differently, an inter-country price index comparison might result in wrong conclusions. Furthermore, the use of CPI and WPI is problematic, as those indices typically include non-tradable goods, which do not comply to the law of one price. For example, a haircut in country A cannot be traded for a haircut in country B, therefore the prices need not be equal. The influence of non-tradable goods on PPP has be studied by several authors, including Strauss (1995) and Dutton (1997). Others, for example Van Dijk and Franses (2006) and Xu (2003), focus on the validity of PPP for the traded goods sector.

Due to the aforementioned problems, the naïve usage and comparison of price indices is risky. In attempt to minimize this risk, I use a price index that is based on traded goods prices only. The inspiration and design for this Trade-weighted Price Index (TPI) is drawn from Xu (2003). To compile this TPI, he uses import and export price deflators which are then weighted by the share of export and import in total trade. The exact construction is explained in detail later in this report. To contrast the results of using TPI with the usage of ordinary CPI, results for both shall be displayed.

2.3.2 The Balassa-Samuelson effect

Closely related to the previously mentioned effect of non-traded goods on PPP is the Balassa-Samuelson effect. This effect is arguably one of the most important arguments for deviations from PPP (Asea and Corden, 1994) and is covered to great extent in the literature. The idea stems from two different articles in 1964, in which Balassa (1964) and Samuelson (1964) independently describe the same phenomenon. The authors observed that productivity growth tends to be larger in the traded goods sector than in the non-traded goods sector. This observation would ultimately lead them to formulate the productivity-biased purchasing power parity hypothesis. Copeland (2008) explains the difference in the productivity growth by noting that the traded goods sector primarily concerns the manufacturing and agricultural sector while the non-traded goods sector concerns the services sector. He goes on to assume that technological innovations (such as automation) are more easily applied in the traded goods sector, which would result in a differential in the productivity. A crucial point is that this differential should be bigger in countries which have a relatively large (technology-sensitive) traded goods sector. Those countries tend to be more industrialized, richer countries. The increase in productivity in the traded goods sector implies that the marginal productivity of labor in that sector has increased relative to that of the non-traded goods sector. Furthermore, the productivity increase should increase wages in the traded goods sector. If we assume that the factor labour can move freely from one sector to the other, this implies that the wage level of both sectors should equalize. Thus, the wage level in the non-traded goods sector will rise so as to match the wage in the traded goods sector. Since productivity has not increased in the non-traded goods sector, the only way to accommodate the wage increase is by raising prices. In the end, the effect is that the country with the positive productivity shock now has a higher price level. This result is in accordance with the observation that countries with higher GDP per capita usually have higher price levels: productivity differentials thus bias PPP.

The literature includes a large body of empirical studies on the Balassa-Samuelson effect. The interested but time constrained reader can be referred to the survey articles by Officer (1976) and Tica and Druzic (2006). Initially, the hypothesis was tested using a cross-section OLS regression in which the real exchange rate was regressed on a proxy of productivity. The hypothesized relationship would be as equation (6):

$$q_t = \alpha + \beta \left(\frac{A^*}{A} \right)_t + \epsilon_t \tag{6}$$

which expresses that the real exchange rate q_t is a function of $\left(\frac{A^*}{A}\right)_t$ which is the ratio of foreign (denoted with *) to domestic productivity. If the Balassa-Samuelson effect exists, we would expect β to be significantly positive. In such a case, relatively higher domestic productivity (an increase in A) leads to higher domestic prices, which results in an appreciation of the real exchange rate (q_t decreases). Using this method, Balassa (1964, 1973) found a significant coefficient, confirming the hypothesis. Some other authors, however, who used this (or a similar) approach did not find similar evidence (for example: De Vries (1968), Clague and Tanzi (1972) and Grunwald and Slazar-Carrillo (1972)). In his survey, Officer (1976) notes that the different results could be due to time- or currency-specific effects in the sample. To avoid such issues, Hsieh (1982) developed a time-series approach and found evidence for the Balassa-Samuelson effect. In a similar manner, but with different proxies for productivity, Edison and Klovland (1987), Marston (1989), Milesi-Ferretti and Micossi (1994) and Rother (2000) also find support for the effect. Yet, using a time-series approach did not imply complete validation of the Balassa-Samuelson hypothesis. Bahmani-Oskooee (1992) finds very mixed long run evidence for the effect, while Asea and Mendoza (1994) fail to find support. In the more recent years, the focus of empirical work on the Balassa-Samuelson effect has shifted towards using cointegration tests. These tests aim to identify a long run relationship between the real exchange rate q_t and proxies for relative productivity $\left(\frac{A^*}{A}\right)_t$. This was first executed with the Engle-Granger cointegration procedure (see section 2.2), using the residuals of equation (6). Later, the use of the Johansen cointegration procedure (see section 2.2) on equations similar to equation (6) gained popularity in the literature. Halikias et al. (1999) find a significant Balassa-Samuelson type cointegration relationship for several European countries, while DeLoach (2001) finds (weak) evidence for a longer sample of OECD countries. In similar vein, Égert (2002) finds generally supportive evidence for the Balassa-Samuelson effect for Eastern European countries. Also noteworthy is the work by Drine and Rault (2003) who show that, although standard time-series cointegration methods reject Balassa-Samuelson effects for Middle Eastern and North African countries, panel cointegration methods strongly support the hypothesis. Overall then, the evidence for the Balassa-Samuelson effect seems mixed, although the previously mentioned survey by Tica and Druzic shows that the largest part of the literature supports the existence of the effect.

2.3.3 Price distortions

Deviations from PPP can also be the result of price distortions which are due to tariffs, quotas, pricing to market, transportation costs and other barriers to trade. For a broader review of some of these issues, one could read Miljkovic (1999). It is simple to imagine how a country's tariff enables sustaining PPP deviations: even without transportation costs an importer now has extra costs to cover. Cassel (1921) already noted the effects of said tariffs on PPP almost a century ago, after observing the beggar-thy-neighbour policies surrounding World War I. 'Pricing to market' concerns the difference in the price elasticity of demand between two countries. This difference can grant suppliers some market power, allowing them to set different prices for every region. This phenomenon is covered in depth by many authors, including Krugman (1986), Dornbusch (1987) and Betts and Devereux (2000). Concerning transportation costs, the literature includes an interesting study by Engel and Rogers (1996), estimating that, when it comes to price dispersion, adding a border between the trading cities has the same effect as adding 75.000 miles between trading cities in the same country. Similarly, Wei and Parsley (1995) find that transportation costs influence deviations from PPP significantly. Chen (2004) finds evidence that deviations from PPP are related to, among others, non-tariff barriers and distances between countries. An interesting model concerning transportation costs is Samuelson's (1954) iceberg model. In this model, part of the product which is being shipped is assumed to 'melt' before it arrives at its destination. Such a formulation allows the transportation costs to be incorporated in a model via the melting. Thus, the importer who wants ten dollars worth of imported goods might need to pay fifteen dollars to get them. Different versions of this iceberg model has been implemented in newer models (see for example: Obstfeld and Rogoff (2000) and Krugman (1998)).

2.4 Nonlinearities

Aside from the iceberg models, deviations from PPP due to transportation costs gave rise to a different class of models, focusing on nonlinearities. For both the iceberg and the nonlinear models, the core idea can be traced back to Heckscher (1916), who noted that deviations from PPP for spatially segregated markets can persist if the costs of arbitrage are higher than the revenue gained from exploiting that arbitrage opportunity. This only allows for small deviations from PPP, since larger deviations increase the benefits of arbitrage beyond the costs, making arbitrage profitable. If this is truly an underlying process for PPP deviations, then we would expect to see nonlinear behavior in the real exchange rate. More specifically, suppose that, in a world without transportation costs, equation (1) holds, so that the real exchange rate is in its PPP-implied equilibrium. If we then impose transportation costs, minor PPP deviations might persist forever. However, the larger the deviation from PPP, the more profitable will arbitrage be and the faster the real exchange rate will return to its PPP-implied value. Thus, the relationship is nonlinear in that larger deviations are corrected faster than smaller deviations. It is important to note that this type of reasoning is not only valid for transportation costs. Dixit (1989) and Obstfeld and Taylor (1997) mention that trade barriers, uncertainty and sunk costs, such as the costs of setting up a distribution network, also play a key role in determining the size of PPP-deviations that the markets allow to exist.

The first empirical works that studied the theoretical implication of nonlinearity assumed that there is a certain fixed band in which deviations from purchasing power parity persist. The aforementioned study by Obstfeld and Taylor refers to this band as the “band of inaction”. Within this band, the real exchange rate is assumed to behave like a nonstationary process. However, if the real exchange rate crosses this threshold, it immediately behaves like a stationary process and therefore tends to return to its parity value. Note that this type of model does *not* allow for growing mean reversion when the deviation from parity increases: it assumes an abrupt change from nonstationarity to mean reversion. All these assumptions were combined with an autoregressive element into a Threshold Autoregressive (TAR) model (Tong, 1990). An easy formulation of this model is given by Taylor (2001). The three separate regimes are given by:

$$q_t = \begin{cases} +c + \rho(q_{t-1} - c) + \epsilon_t & \text{if } q_{t-1} > c \\ q_{t-1} + \epsilon_t & \text{if } c \geq q_{t-1} \geq -c \\ -c + \rho(q_{t-1} + c) + \epsilon_t & \text{if } -c > q_{t-1} \end{cases} \tag{7}$$

where c determines the location of the threshold bands, ρ is the coefficient denoting the degree of mean reversion and ϵ_t is an error term with a standard normal distribution. If the real exchange rate q_t lies between the bands given by c, it behaves like a random walk. Whenever it moves outside of the band, the autoregressive coefficient ρ gradually forces it back between the bands. The model proved successful in empirical application, see for example: Michael et al. (1994).

The “jumpy” behavior of the TAR models could be problematic if the data is characterized by smooth regime changes instead of the abrupt regime changes. Taylor et al. (2001) argue that the usage of time aggregation (taking monthly averages) and commodity aggregation (using price indices) makes for the necessity of a model that allows a smoother transition. The next section will cover the model used in this report that deals with this issue.

3 Methodology

Before any empirical application of exchange rate models takes place, it is useful to get familiar with the used model, the data and the applied methods in this report. First, the specification of the central model of this report, the ESTAR model, is covered. Afterwards, the dataset is introduced and the construction of the series is covered. Third, the used estimation method and the preliminary tests are described. The section ends with a brief introduction to the used dummies and the forecasting methods.

3.1 Model specification

The literature review above ended with the note that the TAR model needs improvements, as it does not allow for a smooth transition between nonstationarity and mean reversion. Allowing for this is important, as the data is characterized by time and price aggregation. The solution contains a function that does not use abrupt regime switching, but rather allows for an area in which the real exchange rate is neither fully nonstationarity nor fully mean reverting. As the real exchange rate moves away from the equilibrium value implied by PPP, it becomes less nonstationary and more mean reverting. A function that incorporates this process was first suggested by Chan and Tong (1986) and is called a Smooth Transition Autoregressive (STAR) model. The most general form of the STAR model is given by:

$$[q_t - \mu] = \sum_{j=1}^{p} \beta_j [q_{t-j} - \mu] + \Phi[\theta, q_{t-d} - \mu] * \sum_{j=1}^{p} \beta_j^* [q_{t-j} - \mu] + \epsilon_t \tag{8}$$

where $[q_t - \mu]$ denotes the deviation of the real exchange rate q_t from the PPP-implied equilibrium value μ and the first term on the right hand side is an ordinary autoregressive component going from period $j = 1$ to p. The second term, $\Phi[\theta, q_{t-d} - \mu]$, is the transition function, which determines the *type* of STAR model. It takes values between zero and one, depending on both the deviation of the real exchange rate from its PPP-equilibrium d periods ago and the value of the parameter θ, which determines the speed of mean reversion. Third, we have another autoregressive component, which is connected to the transition function. Thus, this component only plays a role when $\Phi > 0$. Lastly we have a normally distributed error term, ϵ_t.

3.1.1 The STAR mechanics

The mechanics of this STAR model are as follows. There are 2 separate forces at play. The first can be seen if we set the real exchange rate at period $t-d$ equal to its PPP-implied equilibrium value: $q_{t-d} = \mu$. In that case, by construction, the transition function will become zero: $\Phi = 0$. We are then left with only a simple autoregressive process:

$$[q_t - \mu] = \sum_{j=1}^{p} \beta_j [q_{t-j} - \mu] + \epsilon_t \tag{9}$$

Thus, when PPP holds, the real exchange rate is determined by an AR(p) model. In the literature, it is conventional to set $\beta_1 = 1$ and $p = 1$ so that equation (9) becomes a random walk:

$$[q_t - \mu] = [q_{t-1} - \mu] + \epsilon_t \tag{10}$$

This inner regime of the STAR model is thus governed by a nonstationary process. Now let's move to the outer regime. Assume that d periods ago, the real exchange rate did *not* equal its PPP-implied equilibrium value: $q_{t-d} \neq \mu$. In that case, the transition function Φ can take any value in the range $(0, 1]$ The exact value of the transition function depends on the size of the deviation of the real exchange rate from its PPP-implied equilibrium and on the sensitivity of the model to such deviations. That sensitivity is captured by the parameter θ, that can take any positive value. The transition function then gives a weight to the second autoregressive (AR(p)) process. Assume for a moment that $\Phi = 0.5$ so that the second AR(p) process gets a weight of a half. If we again use the above conventions that $p = 1$ and $\beta_1 = 1$ we are left with:

$$[q_t - \mu] = [q_{t-1} - \mu] + 0.5 * \beta_1^*[q_{t-1} - \mu] + \epsilon_t \tag{11}$$

The last piece of the mechanics of the model is the coefficient β_1^*. For the STAR model to function as theorized, we need this coefficient to take a negative value. This is the case because a negative value for β_1^* combined with the earlier assumption that $\beta_1 = 1$ causes the relationship $\beta_1 + \beta_1^* < 1$ to hold. This relationship is crucial as it assures that the STAR model is mean reverting when deviations from PPP become larger. To see this, note that we can rewrite equation (11) as:

$$[q_t - \mu] = (1 + 0.5 * \beta_1^*)[q_{t-1} - \mu] + \epsilon_t \tag{12}$$

where we need $(1 + 0.5 * \beta_1^*) < 1$ to have a stationary (mean reverting) process. To summarize: when the real exchange rate is equal to the PPP-implied equilibrium, it behaves like a random walk. However, if it deviates from that equilibrium, the transition function together with an autoregressive process assure that the real exchange rate becomes stationary and tends to returns to its equilibrium. The mechanics of the STAR model cause this mean reversion to be stronger when the deviation of the real exchange rate from PPP is bigger.

3.1.2 The PPP-implied equilibrium

Before moving on to the ESTAR model itself, let's consider the role of the PPP-implied equilibrium exchange rate μ. So far, it was assumed to be a fixed number over time, but what is the *value* of this PPP-implied real exchange rate? Absolute purchasing power parity gives us the answer. Start with equation (1) where PPP dictates that the nominal exchange rate exactly equals the price ratio. In that case, the PPP-implied real exchange rate becomes: $R = \frac{SP^*}{P} = 1$. The STAR model used above, however, uses the natural logarithm of the variables. Hence, the *logarithmic* PPP-implied real exchange rate is calculated to equal: $ln(1) = 0$. We thus end up with $\mu = 0$. Using this value, however, is problematic in empirical testing as it presupposes that absolute PPP holds, which is in most cases contrary to what the data suggests.

Fortunately, there is an alternative way to determine the value of μ that is more tailored to the data. This alternative is not based on the strict version of absolute PPP but rather on a weaker version of absolute PPP and on *relative* PPP, for which there is more empirical evidence. In contrast to the analysis before, this alternative PPP-implied equilibrium exchange rate is not fixed, but can vary over time. It therefore gets a time subscript: μ_t. How is a time-varying PPP-implied equilibrium real exchange rate possible? The answer is best explained in the work of Hakkio (1992). First of all, it requires constructing a series of *nominal* exchange rates for which PPP is assumed to hold. After this nominal series is constructed and has had a log transformation, the real exchange rate is obtained via the method shown in equation (4). The first step in constructing the nominal exchange rate series is to assume that equation (1) is replaced by:

$$S = \gamma \frac{P}{P^*} \tag{13}$$

where γ is some constant, with $\gamma \neq 1$. Note that equation (13) is nothing more than a slightly weaker version of absolute PPP. To move from absolute PPP to the relative version of this PPP-equation, we take the first derivative to get the growth rates. As the growth rate of constant γ is zero, it disappears from the equation, so that we are left with the regular relative PPP equation (2). The point here is that for the relative version of PPP to hold, it is *not* required that the strict absolute version of equation (1) holds: the weaker version of equation (13) works just as well as a starting point for relative PPP. The next step is to assume that equation (13) holds at time t. This implies that a weaker version of absolute PPP is assumed to hold at period t. The nominal exchange rate is then denoted as S_t^{PPP}. Then, if relative PPP holds, next period's $(t+1)$ nominal exchange rate is calculated by adjusting that nominal exchange rate for the difference in the inflation rates home and abroad:

$$S_{t+1}^{PPP} = S_t^{PPP}(1 + \pi_t - \pi_t^*) \tag{14}$$

where S_{t+1}^{PPP} is the nominal equilibrium exchange rate at time $t+1$ and π_t and π_t^* are the inflation rates[1]. Hence, using relative PPP, and assuming that a weaker version of absolute PPP holds, next period's PPP-implied nominal exchange rate can be calculated. The same process, also used by Grossmann and McMillan (2010), can be repeated to calculate the nominal exchange rates for period $t+2$, $t+3$ and so on. When this equilibrium nominal exchange rate series has been constructed, the last step is to use equation (4) to get the equilibrium *real* exchange rate μ_t.

The process of constructing the time-varying PPP-implied real exchange rate described above requires setting a base period in which equation (13) holds. Because this decision influences the results, Grossmann et al. (2014) use an approach that uses the entire dataset instead of one base observation. Their "rolling base period" method, described below, is used in this report as well. Consider a dataset that spans the period from January 1980 to December 2013. The first step is to take the average spot exchange rate of the first two years (1980-1981) combined. This average is then used as the starting point S_t^{PPP} for calculating the equilibrium spot exchange rates for every following month (starting with January 1982), using equation (14). This produces one series of equilibrium spot exchange rates. Afterwards, the base period is shifted one year: the average of the second and third year (1981-1982) together is used as the base period. With this new base period, all other equilibrium spot exchange rates are again calculated, which produces another series of equilibrium spot exchange rates. This process goes on until the final base period (2012-2013) is used. Thus, we are left with 33 equilibrium series. The final step is to use these series to get one single equilibrium spot exchange rate for each month. This is achieved by taking the average over all

[1]These inflation rates are simply the shorthand notations of $\frac{P_1 - P_0}{P_0}$ and $\frac{P_1^* - P_0^*}{P_0^*}$ used in equation (2).

equilibrium series for every month.

3.1.3 The ESTAR specification

It is safe to say that most work on STAR models was done by Teräsvirta with several co-authors (see: (Luukkonen et al., 1988), (Granger and Teräsvirta, 1993), (Teräsvirta, 1994), (Eitrheim and Teräsvirta, 1996), (Teräsvirta, 1998) and (Lundbergh et al., 2003)). The STAR model, however, only tells us that the process smoothly returns to its parity, it does not specify *how* it returns. Thus, the model still requires a specification of the transition function Φ. The two most popular functions are the logistic and the exponential function, which give rise to the logistic STAR (LSTAR) and exponential STAR (ESTAR) models:

$$[q_t - \mu_t] = \sum_{j=1}^{p} \beta_j [q_{t-j} - \mu_{t-j}] + \frac{1}{1 + \exp[-\theta[q_{t-d} - \mu_{t-d}]]} * \sum_{j=1}^{p} \beta_j^* [q_{t-j} - \mu_{t-j}] + \epsilon_t \quad (15)$$

$$[q_t - \mu_t] = \sum_{j=1}^{p} \beta_j [q_{t-j} - \mu_{t-j}] + \left(1 - \exp[-\theta[q_{t-d} - \mu_{t-d}]^2]\right) * \sum_{j=1}^{p} \beta_j^* [q_{t-j} - \mu_{t-j}] + \epsilon_t \quad (16)$$

where equation (15) is the LSTAR specification and equation (16) is the ESTAR specification. Both transition functions live on the interval $[0, 1]$, yet the mechanics through which they reach their values are different. Overall, these two models are in many ways similar, but, as Taylor et al. (2001) note, the LSTAR model suggests that the transition function is asymmetric, implying that a negative deviation from the real exchange rate parity is treated differently compared to a positive deviation. This can easily be verified if one fills in the values of the transition function of equation (15). Assume we have $\theta = 1$, $q_{t-d} = 2$ and $\mu_{t-d} = 1$. The resulting value for the transition function is: $\frac{1}{1 + exp[-1[2-1]]} \approx 0.731$. Now take the opposite case, where $q_{t-d} = 1$ and $\mu_{t-d} = 2$ (while we still have $\theta = 1$). The transition function's value then becomes: $\frac{1}{1 + exp[-1[1-2]]} \approx 0.269$. Hence, a positive and negative deviation from equilibrium are treated differently, for which there is no theoretical reason. The ESTAR model does not have the above asymmetry issue, as its transition function works with the *squared* deviation from the parity. This assures that equal positive and negative deviations result in equal values for the transition function. The conclusion is that the (symmetrical) ESTAR model fits the theory better. The data, however, can suggest otherwise. Liew (2004) finds that for some ASEAN members the LSTAR model fits the exchange rate behavior rather well while Teräsvirta and Anderson (1992) successfully fit an LSTAR model to some industrialized countries. Yet, the literature also holds evidence of ESTAR-type nonlinearities, see for example Michael et al. (1997), Taylor et al. (2001) and Grossmann and McMillan (2010).

Because of the LSTAR asymmetry discussed above, this report focuses on the ESTAR model. Equation (16) is however not the specification that is estimated. Three conventions are used to simplify the specification. The first is the convention that the inner regime is characterized by a random walk, so that $\beta_1 = 1$. The second simplification concerns the value of β_1^*. In one of the most influential works on the ESTAR model, Taylor et al. (2001) set $\beta_1^* = -1$ *a priori* without specifying why this assumption is made. Buncic (2007) criticizes this assumption as the value of β_1^* has an impact on the size of the transition speed θ. He goes on to find that larger negative values of β_1^* are associated with lower values of θ. The precise statistical relationship between β_1^* and θ is outside the scope of this report, but it is easy to verify that having a small value for θ undermines the nonlinear regime. Consider the implication of having $\theta \approx 0.3$, as was a finding in Taylor et al. (2001). Since the deviation of the logged real exchange rate from its parity value is generally somewhere in the range $[-1, 1]$, the transition function will then only take a value between 0 and 0.26[2]. Hence, low values of θ imply that only a small part of the range of the transition function range (which is $[0, 1]$) is being used in the entire model. As Buncic (2007) notes, if none of the observations ever gets the full weight of 1 from the transition function, this makes the notion of a regime

[2]If $\theta \approx 0.3$ and $q_{t-1} - \mu_{t-1} = 1$, we have that $\Phi = 1 - \exp[-0.3 * 1^2] \approx 0.26$.

questionable. To avoid these issues, Buncic suggests using $\beta_1^* = -0.1$. Using such a relatively low value for β_1^* is supported by the findings of Grossmann and McMillan (2010), who find similar low values when estimating that parameter. In light of their work, one might ask why β_1^* is not just freely estimated in this report. The answer is that freely estimating β_1^* in this dataset results in non-convergence and local minimums which are not in line with the theoretical framework. Therefore, the second convention used here is that $\beta_1^* = -0.1$.

The third convention is the specification that $p = 1$ and $d = 1$. This implies that there are no higher order autoregressive processes and that the delay between the deviation from the PPP-implied equilibrium and the correction to the current real exchange rate is just one period. This convention is used broadly in the literature and keeps the specification parsimonious. After applying these conventions, the ESTAR specification is:

$$[q_t - \mu_t] = [q_{t-1} - \mu_{t-1}] + \left(1 - \exp[-\theta[q_{t-1} - \mu_{t-1}]^2]\right) * -0.1[q_{t-1} - \mu_{t-1}] + \epsilon_t \quad (17)$$

where θ is the crucial coefficient to be estimated. Note that the use of $\Phi = 1-\exp[-...]$, ascertains that the value for the transition function will always be in the range $[0, 1]$. A further simplification is to rewrite equation (17) by subtracting $[q_{t-1} - \mu_{t-1}]$ from each side to get:

$$\Delta[q_t - \mu_t] = \left(1 - \exp[-\theta[q_{t-1} - \mu_{t-1}]^2]\right) * -0.1[q_{t-1} - \mu_{t-1}] + \epsilon_t \quad (18)$$

where $\Delta[q_t - \mu_t]$ denotes the first difference of the (log) real exchange rate deviation from the PPP-implied equilibrium. This first difference specification of the ESTAR model has the side benefit of avoiding spurious regressions that could occur when estimating equation (17) due to nonstationarity. The last step in this specification process is to add an autoregressive component in equation (18). This extra term is unprecedented in the literature; I add it so as to avoid serial correlation in the residuals of the ESTAR specification[3]. Only the first lag of the difference is added, as any extra lagged differences for any currency pair had no significant value added when estimated. When adding the autoregressive term, the final specification is given by:

$$\Delta[q_t - \mu_t] = \left(1 - \exp[-\theta[q_{t-1} - \mu_{t-1}]^2]\right) * -0.1[q_{t-1} - \mu_{t-1}] + \lambda(\Delta[q_{t-1} - \mu_{t-1}]) + \epsilon_t \quad (19)$$

where λ is the coefficient of the autoregressive component. Equation (19) is the specification that is empirically tested in this report, where θ and λ are estimated.

3.2 Data and series construction

The used data covers the period of January 1980 up to December 2013 at a monthly frequency (408 observations), for the US dollar (the base currency), the Japanese yen, the euro (and Dutch guilder, as the Netherlands is used as representative country), the Canadian dollar and the Singapore dollar. These currencies have the advantage of having the necessary data, and having a more or less floating exchange rate during the sample. Without a floating exchange rate, there is no sound theoretical argument why PPP should hold at all, thus having a floating exchange rate is of importance for this report. Nearly all data was retrieved online from the International Monetary Fund's International Financial Statistics database. The last few export and import deflator observations for the TPI construction for the Netherlands were retrieved from the Dutch CBS StatLine. Unless specified otherwise, the data concerns monthly averages. All index values have 2005 as the base year. Note that the exchange rate is defined as the number of domestic currency units per unit foreign currency. Here, the US is the base country, which makes the US dollar take the role of being the 'foreign' currency. The dollar-yen exchange rate is then specified as the amount of yen received for one US dollar.

[3]Readers familiar with the literature on the Dickey Fuller test might note that this extra autoregressive term is comparable to the augmentation in the Augmented Dickey Fuller test, which was also my inspiration for adding this term

A first important step is to construct the real exchange rate series for the sample period and currencies. The real exchange rate is constructed as specified in equation (4). Thus, it requires three series: the nominal exchange rate s_t, the domestic price level p_t and the foreign price level p_t^*. Two different real exchange rate series are built; the first one using Consumer Price Index data and second one using the Trade Price Index. These two price indices are combined with the same nominal exchange rate series to get the two real exchange rate series. The construction of the CPI-based real exchange rate is fairly straightforward: all required series are readily available as monthly averages and can be inserted in equation (4) after taking the natural logarithm. The TPI-based exchange rate series, however, require some assembling and construction. The TPI for country i at time t is built using Xu's (2003) approach:

$$TPI_{i,t} = \frac{V_{i,t}^{EX}}{V_{i,t}^{EX} + V_{i,t}^{IM}} * P_{i,t}^{EX} + \frac{V_{i,t}^{IM}}{V_{i,t}^{EX} + V_{i,t}^{IM}} * P_{i,t}^{IM} \tag{20}$$

where $V_{i,t}^{EX}$ and $V_{i,t}^{IM}$ are the total export and import values in period t respectively (measured in US dollars), and $P_{i,t}^{EX}$ and $P_{i,t}^{IM}$ are the respective export and import unit deflators. Thus, the TPI is constructed as a weighted average of export and import deflators, with the weights being the respective share of exports and imports in total trade value. Note that the TPI series also uses 2005 as the base year.

3.3 Estimation methods and preliminary tests

With these real exchange rates constructed, some preliminary unit root tests are performed. Testing for unit roots is done using the Augmented Dickey Fuller (ADF) test, where the possible inclusion of an intercept and trend coefficient is decided on through visual inspection of the series graphs. The lag structure for the ADF regression is decided on by the Schwarz Information Criterion, allowing for a maximum of 36 lags.

Before the coefficients of the ESTAR model are calculated, the real exchange rate series are checked for nonlinearity with a test devised by Kapetanios et al. (2003)

which is denoted here as the KSS test. It is important to note that this test uses the null hypothesis of *nonstationarity* while the alternative is *nonlinearity.* Originally, the KSS test was designed as an alternative for the Augmented Dickey Fuller unit root test, considering that the results of the ADF test might lack power if the real underlying process is nonlinear. To derive the test, consider the ESTAR specification of equation (18), but drop the assumption that $\beta_1^* = -0.1$. Hence, we put β_1^* back in the specification, instead of -0.1. The transition speed coefficient θ can take any positive value, where larger values denote faster mean reversion. To test whether the real exchange rate is characterized by nonlinearities, one would want to test the null hypothesis $H_0 : \theta = 0$ against the alternative that $H_a : \theta > 0$. Notice that under the null hypothesis equation (18) is a random walk. Directly testing the null hypothesis is not possible: inserting $\theta = 0$ in equation (18) leads to β_1^* being multiplied with 0. Hence, for any value of β_1^*, the results would be the same: β_1^* can thus not be identified under the null hypothesis. The solution for this problem is using a first-order Taylor approximation. The (fully written out) first-order Taylor approximation[4] for equation (18) around $H_0 : \theta = 0$ is:

$$\Delta[q_t - \mu_t] = -[q_{t-1} - \mu_{t-1}]^2 * -\beta_1^*[q_{t-1} - \mu_{t-1}] \exp[0] * \theta + \epsilon_t \tag{21}$$

which is easily simplified to:

$$\Delta[q_t - \mu_t] = \delta[q_{t-1} - \mu_{t-1}]^3 + \epsilon_t \tag{22}$$

where $\delta = \theta\beta_1^*$. This coefficient δ can then be estimated using OLS, with $H_0 : \delta = 0$ and $H_a : \delta < 0$.[5] If δ is significantly below zero, the null hypothesis of nonstationarity is rejected. Just like the Augmented Dickey Fuller test, the KSS test allows for augmentations through the addition of lags so as to get serial uncorrelated errors. This

[4]The first-order Taylor approximation formula for the function $f(x)$ is: $f(x_0) + f'(x_0)(x - x_0)$. In our case, $x = \theta$ and $x_0 = 0$.

[5]Recall that under the alternative $\theta > 0$ while $\beta_1^* < 0$, so that $\delta = \theta\beta_1^* < 0$

augmented KSS test is given by:

$$\Delta[q_t - \mu_t] = \delta[q_{t-1} - \mu_{t-1}]^3 + \sum_{j=1}^{p} \phi_j \Delta[q_{t-j} - \mu_{t-j}] + \epsilon_t \tag{23}$$

where p is the number of included lags. Equation (23) is the specification used in this report, where p is decided on through inspection of the explanatory power of the extra lag when it is added to the specification. In one of the works on the KSS test, Kapetanios et al. (2003) provide the critical values for the t-Statistics of the test, as it does not follow the asymptotic standard normal distribution. The reported critical values, however, correspond to numbers of observations which are not close to the number of observations used in this report.

Therefore, I use Monte Carlo simulation techniques to generate the t-Statistic probability values that do correspond with the used observation number. This implies that a distribution for t-Statistics of δ is simulated while the random walk is the true data generating process. The first step is to simulate random walk data and estimate a KSS specification (that is; equation (23)) on that data. Each random walk simulation consists of 508 observations of which the first 100 are dropped so as to match this report's observation count of 408. Thus, over these remaining 408 random walk observations an KSS equation is estimated and the corresponding t-Statistic of δ is stored. This process is repeated for 5000 simulations. Afterwards, the actual t-Statistics are compared with these 5000 simulated t-Statistics. The percentage of simulated t-Statistics that lie above the (absolute) value of the actual t-Statistic is then used as the simulated probability belonging to δ. These probabilities can be found in table A.2 in the Appendix and are used to determine the significance of the estimated δ coefficients.

The estimation of the coefficients λ and θ of the ESTAR specification of equation (19) is performed using Nonlinear Least Squares (NLS). In contrast to OLS , NLS has no closed form solution so the Nonlinear Least Squares estimator must be derived using numerical iteration methods. For this method to give consistent results, two aspects are of importance. First of all, there must be a unique global minimum for the least squares estimator so that the model is identified. Furthermore, the starting values of the parameters should be as close as possible to the true values so as to avoid hitting local minimums. The starting values used in this report are inspired by the work of Taylor et al. (2001) and Grossmann and McMillan (2010). For the Netherlands, the extra autoregressive component with the coefficient λ showed no significant explanatory power and was thus left out of the estimation.

The evaluation of the estimated θ coefficients creates difficulties similar to the case of the KSS test. Again, we have that under $H_0 : \theta = 0$ the specification of equation (19) collapses to become a random walk. Hence, the probabilities associated with the t-Statistic are not valid as the t-Statistic does not follow a t-distribution under the random walk. The solution is the same as in the case of the KSS test: the correct probability values can be obtained through simulations. This approach is similar to the one used by Taylor et al. (2001). Just like described before in the case of the KSS test, a distribution for the t-Statistics is simulated using Monte Carlo methods, for which the random walk is the underlying process. Again, random walk data covering 508 observations is generated, after which the first 100 observations are dropped so as to get 408 observations. Then, an ESTAR equation like equation (19) is estimated using those observations. This allows the resulting t-Statistic of θ to be obtained. Following this procedure, 5000 simulated t-Statistics are stored. Just as before, the actual t-Statistics are compared with these 5000 simulated t-Statistics and the percentage of simulated t-Statistics that lie above the (absolute) value of the actual t-Statistic is used as the simulated probability belonging to θ. These probabilities can be found in table A.3 in the Appendix and are used to determine the significance of the estimated θ coefficients.

3.4 Dummies and forecasts

The ESTAR coefficients are estimated multiple times for all four country pairs: once with the CPI- and TPI-based real exchange rates, but also with an extra added dummy variable that captures the effect of the recent financial downturn of the years 2008 to 2010. Why add a dummy? One would expect the costs of uncertainty to increase in times of economic crisis, which would result in less arbitrage. If this is truly the case, then the extra coefficient should display a significant impact during the mentioned years. This dummy variable is attached to the transition function in the ESTAR model. Hence, the ESTAR specification with dummy variable is:

$$\begin{aligned}\Delta[q_t - \mu_t] = &\big(1 - \exp[-\theta[q_{t-1} - \mu_{t-1}]^2]\big) * -0.1[q_{t-1} - \mu_{t-1}] \\ &+ D * \big(1 - \exp[-\theta[q_{t-1} - \mu_{t-1}]^2]\big) * \beta_D^*[q_{t-1} - \mu_{t-1}] \\ &+ \lambda(\Delta[q_{t-1} - \mu_{t-1}]) + \epsilon_t \qquad (24)\end{aligned}$$

where D is the dummy variable and β_D^* is the coefficient associated with that dummy component. If there is less nonlinear mean reversion during years of financial downturn, we would expect to find a positive β_D^*. In order to ease convergence in the estimation, it is assumed that the transition speed remains unaffected; that is, θ is invariant to the dummy variable. In reality, this assumption might not hold: the transition speed could very well be lower in times of crisis. This assumption is therefore the downside of this dummy specification.

The final part of this report forms an attempt to forecast the real exchange rate using the ESTAR model. The forecasts are based on a slightly modified version of equation (17):

$$\begin{aligned}[\hat{q}_t - \mu_t] = [q_{t-1} - \mu_{t-1}] + &\big(1 - \exp[-\theta[q_{t-1} - \mu_{t-1}]^2]\big) * -0.1[q_{t-1} - \mu_{t-1}] \\ &+ \lambda(\Delta[q_{t-1} - \mu_{t-1}]) + \epsilon_t \qquad (25)\end{aligned}$$

where $\hat{q}_t$ denotes the forecast of the real exchange rate at time t and the autoregressive component with the coefficient λ was added. After the forecasts of the deviations from the equilibrium PPP are retrieved from equation (25), it is simple to retrieve the forecasted real exchange rates; just add μ_t to get: $[\hat{q}_t - \mu_t] + \mu_t = \hat{q}_t$.

The forecasting is executed via a "recursive regression". This means that the observations from 1980M1 until 1999M12 are used to estimate the model coefficients, which are then used to forecast the real exchange rate for 2000M1. Using the *actual* value for 2000M1, the model coefficients are then again re-estimated and using that updated model, the real exchange rate for 2000M2 is forecasted, and so on. In short, the model is updated every period before forecasting the next period. This process continues until the end of the sample, so the last forecast (2013M12) is based on a model estimated with actual data from 1980M1 to 2013M11. The accuracy of these forecasts is compared to that of a random walk model, which is used as a benchmark since it embodies the concept of unpredictability. Moreover, the performance is also compared to that of a simple autoregressive model with one lag (AR(1) model), which is used as a benchmark naïve model in some empirical works. The performance is compared using the Root Mean Squared Error (RMSE) which is given by:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(q_t - \hat{q}_t)^2} \tag{26}$$

where $q_t - \hat{q}_t$ denotes the gap between the actual and the forecasted value for the real exchange rate at time t. The lower the RMSE, the better the forecast performance.

4 Results

In this section, the results of the methods described above are presented. The results of the preliminary unit root test are presented first, followed by the KSS test. The ESTAR coefficient estimation is covered thereafter. It concludes with the addition of a dummy variable and the forecasting attempt.

4.1 Preliminary tests

After the log transformation of the spot nominal exchange rate and the CPI and TPI series, the series are tested for nonstationary behavior using the ADF test. The results are found in table A.1 in the Appendix. As is clear from table A.1, for most series the null hypothesis of a unit root cannot be rejected. The exceptions are the CPI of Canada, the Netherlands and the US. The unit root hypothesis would only be rejected for the Dutch spot exchange rate at the 10 percent significance level. All nonstationary series are integrated in the order 1, indicating that the first difference of these series is stationary.

To get an idea of how the (nominal) PPP-implied equilibria take shape compared to the actual nominal exchange rate, figures A.1 to A.4 in the Appendix show the relevant series for all currency pairs. Both PPP-implied equilibria generally follow the trend of the actual nominal exchange rate. The CPI-based equilibrium shows less volatility then the TPI-based equilibrium. This implies that the measured import and export deflators tend to vary more than the CPI price level which takes into account nontradable goods.

Table 1 reports the results for the KSS test. All δ coefficient values are negative, which is in accordance with the ESTAR specification. However, only two out of eight coefficients are significantly below zero. For the Japanese CPI and the Singapore TPI version of the real exchange rate deviation from PPP, the null of nonstationarity is rejected. For the Netherlands, the difference between the coefficients for the CPI and TPI versions is negligible, while the difference is the largest for Canada. All in all, this gives no real support for the usage of TPI in ESTAR estimation, as CPI seems to perform just as well. Furthermore, these results could be interpreted as a warning that the ESTAR model will not fit all exchange rate series, as most series seem to lack ESTAR nonlinearity.

Table 1: Preliminary test: KSS test

Country	Estimated δ	t-Statistic	Number of lags (p) included
CPI			
Canada	-0.30	-1.74	1
Japan	-0.16	-2.32*	1
Netherlands	-0.01	-0.40	0
Singapore	-0.14	-0.79	1
TPI			
Canada	-0.11	-1.55	1
Japan	-0.04	-1.97	1
Netherlands	-0.01	-1.04	1
Singapore	-0.26	-2.15*	1

Results for the KSS test as formulated in equation (23) with $H_0 : \delta = 0$ (nonstationarity) and $H_a : \delta < 0$ (nonlinearity). Coefficients that are significantly lower than 0 at the 10, 5 or 1 percent significance level are marked with *,** and *** respectively. The significance levels for δ were obtained via Monte Carlo simulations, see table A.2 in the Appendix. The number of lags with which the specification was augmented is shown in the last column.

4.2 ESTAR estimation

The estimated coefficients of the ESTAR model in table 2 are in line with the KSS test results. For three out of eight currency pairs the coefficient θ is significant at the 10 or 5 percent level. All coefficients show the correct theoretical sign. The sizes of the estimated θ coefficients seem acceptable, and are comparable with the findings of Grossmann and McMillan (2010). In contrast with their findings, however, is the result that TPI-based ESTAR models only perform marginally better than their CPI counterparts.

Table 2: ESTAR estimation

Country	Estimated θ	Estimated λ
CPI		
Canada	3.19	0.23^{***}
	[2.00]	[0.05]
Japan	1.71^{**}	0.31^{***}
	[0.87]	[0.05]
Netherlands	0.06	-
	[0.14]	-
Singapore	1.45	0.22^{***}
	[1.91]	[0.05]
TPI		
Canada	1.12	0.12^{**}
	[0.80]	[0.05]
Japan	0.38^{*}	0.24^{***}
	[0.21]	[0.05]
Netherlands	0.10	0.14^{***}
	[0.10]	[0.05]
Singapore	3.08^{*}	0.21^{***}
	[1.62]	[0.05]

The estimated ESTAR coefficients, from the specification in equation (19) estimated using Nonlinear Least Squares. The standard errors of the coefficients are displayed in parentheses below each coefficient. Coefficients that differ significantly from 0 at the 10, 5 or 1 percent significance level are marked with *,** and *** respectively. The significance levels for θ were obtained via Monte Carlo simulations, see table A.3 in the Appendix.

As was predicted by the KSS test, the ESTAR model does not fit the Dutch real exchange rate well. Perhaps this is due to a structural break in the Dutch nominal exchange rate data in January 1999, when the Dutch guilder was replaced for the euro. The addition of the autoregressive component appears fruitful, as λ is found to be significant for every case except the Dutch CPI-based ESTAR model. Although the addition of this AR component is unprecedented in the literature, it seems a recommendable strategy based on these results.

4.3 Dummy analysis

Table 3 reports the output from the NLS estimation of equation (24). The object of interest is the coefficient β_D^*. If international goods traders are more reluctant to seize arbitrage opportunities created by deviations from the law of one price in times of financial turmoil, we would expect to find a positive value for this coefficient. A positive coefficient would move the ESTAR model away from nonlinear mean reversion and towards unit root behavior. Table 3 does clearly not support this theory, as many of the estimated coefficients are negative and all of them are insignificant. Moreover, the estimated coefficients of the CPI-based ESTAR model for Singapore and the TPI-based model for Japan are (insignificantly) positive but larger than 0.1. Hence, implementing them in the full ESTAR model under our assumptions that $\beta_1 = 1$ and $\beta_1^* = -0.1$ would imply that the deviation from the PPP-implied equilibrium is characterized by explosive behavior, as $\beta_1 + \beta_1^* + \beta_D^* > 1$. As none of the coefficients are significant for either the CPI-based or the TPI-based ESTAR model, the use of a trade-weighted price index has no extra value in the dummy analysis.

The inclusion of the dummy variable has virtually no impact on the estimated value of θ, which also supports the view that a crisis has little impact on nonlinear behavior of the real exchange rate. The three existing ESTAR relationships from table 2 all survive the addition of the dummy. Note that the significance of this transition speed coefficient is again evaluated with the same simulated t-Statistics distribution used in table 2. Unsurprisingly, the added AR components keep their significant explanatory power in the ESTAR model when the dummy is included.

Country	Estimated θ	Estimated β_D^*	Estimated λ
CPI			
Canada	2.73	-0.36	0.22***
	[2.00]	[0.46]	[0.05]
Japan	1.69**	-5.52	0.31***
	[0.87]	[4.21]	[0.05]
Netherlands	0.01	-0.76	-
	[0.18]	[10.15]	-
Singapore	1.48	0.67	0.22***
	[1.91]	[2.26]	[0.05]
TPI			
Canada	0.90	-0.07	0.12**
	[0.93]	[0.23]	[0.05]
Japan	0.40*	0.15	0.24***
	[0.22]	[0.32]	[0.05]
Netherlands	0.07	-0.11	0.14***
	[0.13]	[0.44]	[0.05]
Singapore	3.03*	-4.04	0.21***
	[1.61]	[3.16]	[0.05]

Table 3: ESTAR crisis dummy estimation

The estimated ESTAR coefficients, from the specification in equation (24) estimated using Nonlinear Least Squares. The standard errors of the coefficients are displayed in parentheses below each coefficient. Coefficients that differ significantly from 0 at the 10, 5 or 1 percent significance level are marked with *,** and *** respectively. The significance levels for θ were obtained via Monte Carlo simulations, see table A.3 in the Appendix.

4.4 Forecasting the real exchange rate

The final part of this report consists of a forecasting attempt using the ESTAR model. All recursive forecasts were made for the period of January 2000 to December 2013.

The resulting forecast graph for the CPI-based ESTAR model for the Japanese yen is displayed in figure 1, where the dotted line is the forecast and the continuous line is the actual real exchange rate. Note that the forecast always seems to lag one period behind the actual real exchange rate. This is caused by the assumption that $\beta_1 = 1$ in equation (25).

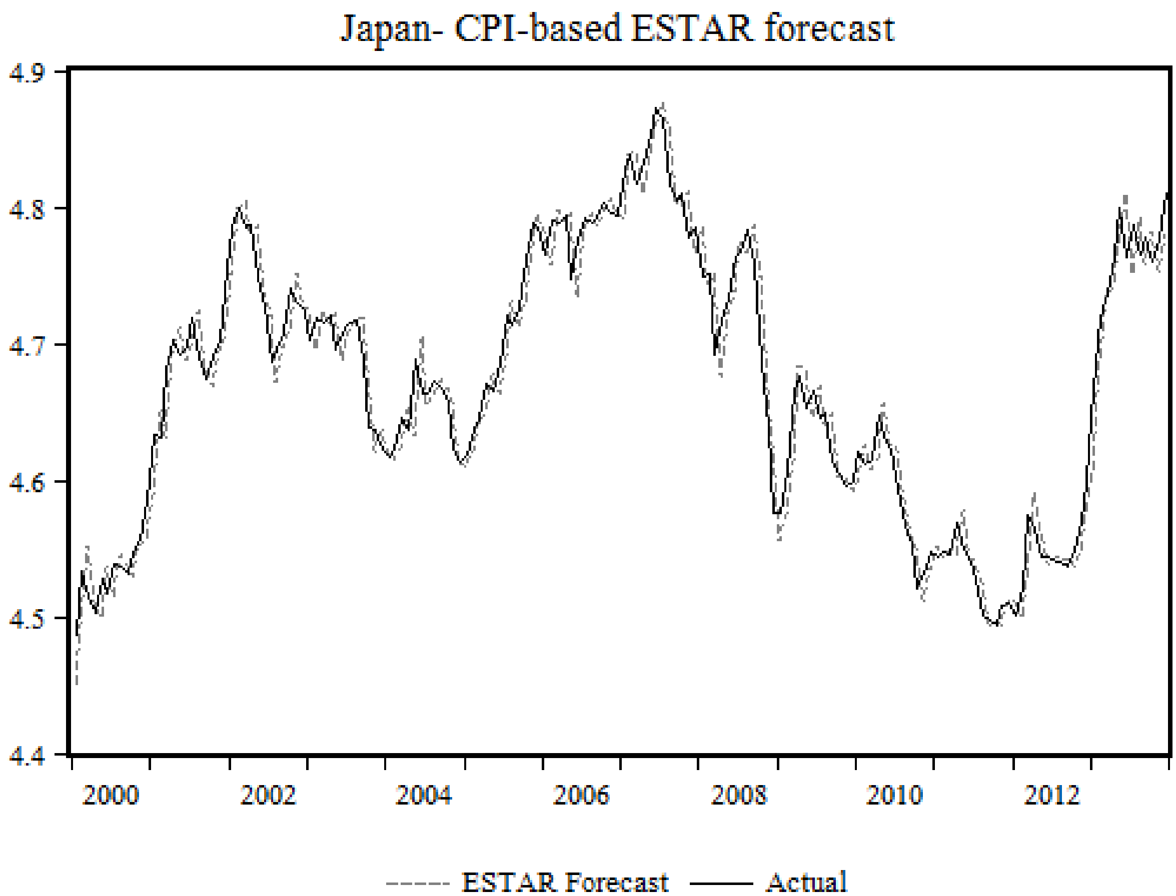

Figure 1: The actual Japanese yen real exchange rate versus the recursive CPI-based ESTAR forecast.

Furthermore, note that the forecast mostly seems to deviate from actual values at the turning point of downward or upward spikes. This implies that the ESTAR model predicts bigger and longer deviations from the PPP-implied equilibrium than is actually the case. Part of these peaks is explained by the extra autoregressive component (λ) in equation (25). That component uses last period's increase or decrease in the deviation from PPP-implied equilibrium and through the positive value of λ adds the expectation that this increase or decrease will persist in the next period. If the true value then shows a turning point, this creates the deviation of the forecast from the actual value. To see the impact, compare figure 1 of Japan with figure A.6 for the euro in the Appendix, in which this extra AR(1) component was left out in the forecast. This does however *not* imply that adding the AR term is a bad idea, as it might add precision in times of continued increasing or decreasing deviations from the PPP-implied equilibrium.

Table 4: Forecast root mean squared errors

Country	ESTAR model	Random walk model	AR(1) model
CPI			
Canada	0.01884	0.01937	0.01938
Japan	0.02386	0.02448	0.02433
Netherlands	0.02436	0.02400	0.02455
Singapore	0.01341	0.01374	0.01368
TPI			
Canada	0.02715	0.02767	0.02730
Japan	0.03089	0.03151	0.03115
Netherlands	0.04333	0.04482	0.04464
Singapore	0.01858	0.01979	0.01924

The root mean squared error (RMSE) as calculated using equation (26) for the ESTAR model, a random walk model and an autoregressive model with a single term (AR(1)). A lower value indicates better forecasting performance.

Overall, it is hard to tell from the figures whether the forecasting performance of the ESTAR model is decent or not. To assess this, table 4 shows the root mean squared error (RMSE) for each model. Lower values denote smaller deviations of the forecast from the actual real exchange rate and thus better performance. The benchmark with which to compare the ESTAR forecasts is the random walk model, which is nothing more than saying that the forecast for next period is the actual value of this period.

For all but one case (the CPI-based ESTAR model for the Netherlands), the ESTAR model makes better forecasts than the random walk. Interestingly, the differences between the random walk RMSEs and the TPI-based ESTAR model RMSEs are larger than those between the CPI-based ESTAR model and the random walk. This implies that the TPI ESTAR models have better added forecasting precision over the random walk than the CPI ESTAR models do. This could support the view that TPI is a better starting point for exchange rate forecasts than CPI, which is in line with the findings of Grossmann and McMillan (2010).

The third column of table 4 gives the RSME for an AR(1) model forecast. This model can generally be seen as the most basic modelling attempt for forecasting besides the random walk. In this case, the AR(1) model only performs marginally better than the random walk model in six out of eight currency pairs. This is due to β_1 taking a value very close to unity in the AR(1) model: $(q_t - \mu_t) = \beta_1(q_{t-1} - \mu_{t-1})$. This makes the AR(1) model behave almost completely in sync with the random walk. Although this implies that the AR(1) model is obsolete in our forecasting exercise, it does confirm that the assumption that $\beta_1 = 1$ made in the specification of the ESTAR model is quite accurate.

5 Concluding remarks

In this report the central theme was the Exponential Smooth Transition Autoregressive (ESTAR) model which describes how the real exchange rate behaves when it deviates from purchasing power parity. Although the real exchange rate can behave like a random walk when it is near its PPP-implied equilibrium, it should show nonlinear mean reversion when it deviates too much from that equilibrium. In this report, this concept was combined with empirical and theoretical literature on PPP and real exchange rate models from which several insights and methods were obtained and used. In this section, the results of these methods and applications are briefly summarized.

When it comes to behavior of exchange rates, it seems there is still enough to be learned. Traces of ESTAR nonlinearity could only be found in two out of eight cases, and the deviations from the PPP-implied equilibrium seem to be mostly governed by a unit root, as was apparent from the KSS test. It is important to note that the empirical literature showed that results of the ESTAR model are influenced by the assumptions made about the model parameters. More research into the correctness of those assumptions could help understand the misfits of the model. However, under the assumptions made here, in the cases in which significant ESTAR nonlinearity *was* in fact found, the transition speed coefficient took on acceptable values. This could imply that the assumptions in this report are a step in the right direction.

The distinction between TPI and CPI made in this report is based on the possible distorting effect on PPP of the inclusion of nontradable goods in the CPI. As the Trade-weighted Price Index only includes traded goods prices, TPI is expected to be a better measure of price difference and movement between countries. When it comes to the ESTAR model, this trait of the TPI is not readily apparent. Throughout the ESTAR estimation, TPI accounted for two significant ESTAR cases (Japan and Singapore) while CPI accounted for one (Japan). Although this is evidence in favor of the TPI-based ESTAR model, the difference is marginal. Stronger evidence came from the forecasting attempt, where the TPI-based ESTAR models made consistently better forecasts than the CPI-based models. Overall, it does not seem that TPI has proven its worth, although it would be interesting to see whether the better forecast performance holds for other models and currencies.

The addition of a dummy variable for the financial crisis to the ESTAR model had very clear-cut results. There is no support for the view that nonlinear behavior of the real exchange rate works differently in times of recession. It should however be noted that this evidence is based on the assumption that the nonlinear transition speed does not change in times of crisis, which can be criticized for being unrealistic. Dropping this assumption, on the other hand, requires estimation techniques that are more consistent, for estimations based on Nonlinear Least Squares are to susceptible to local minimums. The alternative is that traders simply do not change their behavior in times of financial turmoil. That alternative, however, seems hard to believe in the light of historical global trade volume figures.

When it comes to forecasting, the ESTAR models seem to do a reasonable job. For seven out of eight cases, the ESTAR models gave more accurate predictions than a naïve random walk model or an AR(1) model. Interestingly, this result holds also for the cases in which no significant ESTAR nonlinearity was found. It seems likely that the added autoregressive component to the ESTAR model is responsible for some of that performance. This can be derived from the fact that the only case lacking such a component (the Dutch CPI-based ESTAR model) performed worse than the random walk. Peculiarly enough, that same autoregressive component makes for the spiked misfits that can be observed in the forecast figures. Hence, there are two sides to that story.

There remains work to do when it comes to modelling exchange rate behavior. This report provides weak support for the nonlinearity of that behavior, but the ESTAR model does not fit every case and relies on a broad set of assumptions for the cases it does support. Further research could focus on those assumptions and on different types of nonlinearity. After all, the STAR model leaves room for other types of transition functions that perhaps fit the data better. It would be interesting to test those alternative STAR models with the insights and applications used in this report.

References

Adler, M. and Lehman, B. (1983). Deviations from purchasing power parity in the long run. *Journal of Finance*, 39:1471–1487.

Asea, P. K. and Corden, W. M. (1994). The Balassa–Samuelson model: An overview. *Review of International Economics*, 2(3):191–200.

Asea, P. K. and Mendoza, E. (1994). The Balassa-Samuelson model: a general-equilibrium appraisal. *Review of International Economics*, 2(3):244–267.

Bahmani-Oskooee, M. (1992). A time-series approach to test the productivity bias hypothesis in purchasing power parity. *Kyklos*, 45(2):227–236.

Balassa, B. (1964). The purchasing-power parity doctrine: A reappraisal. *The Journal of Political Economy*, 72(6):584–596.

——— (1973). Just how misleading are official exchange rate conversions? A comment. *The Economic Journal*, 83(332):1258–1267.

Betts, C. and Devereux, M. B. (2000). Exchange rate dynamics in a model of pricing-to-market. *Journal of International Economics*, 50(1):215–244.

Buncic, D. (2007). Issues with ESTAR models for real exchange rate modelling. Unpublished Manuscript.

Cassel, G. (1921). *The World's Monetary Problems*. E.P. Dutton, New York.

Chan, K. S. and Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis*, 7(3):179–190.

Chen, N. (2004). The behaviour of relative prices in the European Union: A sectoral analysis. *European Economic Review*, 48(6):1257–1286.

Clague, C. K. and Tanzi, V. (1972). Human capital, natural resources and the purchasing-power parity doctrine: Some empirical results. *Economia Internazionale*, 25(1):3–18.

Copeland, L. S. (2008). *Exchange Rates and International Finance.* Financial Times Prentice Hall, Harlow [etc.].

Corbae, D. and Ouliaris, S. (1988). Cointegration and tests of purchasing power parity. *The Review of Economics and Statistics*, 70(3):508–511.

De Vries, M. G. (1968). Exchange depreciation in developing countries. *Staff Papers*, 15(3):560–578.

DeLoach, S. B. (2001). More evidence in favor of the Balassa–Samuelson hypothesis. *Review of International Economics*, 9(2):336–342.

Dixit, A. (1989). Hysteresis, import penetration, and exchange rate pass-through. *The Quarterly Journal of Economics*, 104(2):205–228.

Dornbusch, R. (1987). Exchange rates and prices. *The American Economic Review*, 77(1):93–106.

Drine, I. and Rault, C. (2003). A re-examination of the Balassa–Samuelson hypothesis using recent panel data unit-root and cointegration tests: evidence from MENA countries. *African Development Review*, 15(2/3):106–125.

Dutton, M. (1997). Cointegration tests of purchasing power parity: The impact of non-traded goods. *Journal of International Money and Finance*, 16(3):433–444.

Edison, H. J. (1987). Purchasing power parity in the long run: A test of the dollar/pound exchange rate. *Journal of Money, Credit and Banking*, 19:376–387.

Edison, H. J. and Klovland, J. (1987). A quantitative reassessment of the purchasing power parity hypothesis: Evidence from norway and the United Kingdom. *Journal of Applied Econometrics*, 2(4):309–333.

Égert, B. (2002). Estimating the impact of the Balassa–Samuelson effect on inflation and the real exchange rate during the transition. *Economic Systems*, 26(1):1–16.

Einzig, P. (1970). *The History of Foreign Exchange.* Macmillan, London.

Eitrheim, Ø. and Teräsvirta, T. (1996). Testing the adequacy of smooth transition autoregressive models. *Journal of Econometrics*, 74(1):59–75.

Enders, W. (1988). ARIMA and cointegration tests of PPP under fixed and flexible exchange rate regimes. *The Review of Economics and Statistics*, 70(3):504–508.

Engel, C. and Rogers, J. H. (1996). How wide is the border? *The American Economic Review*, 86(5):1112–1125.

Engle, R. F. and Granger, C. J. W. (1991). Cointegration and error correction: Representation, estimation, and testing. *Econometrica*, 55:251–276.

Fisher, E. O. (1991). Testing purchasing power parity under the null hypothesis of cointegration. *Economic Journal*, 101(409):1476.

Frankel, J. (1986). International capital mobility and crowding out in the U.S. economy: Imperfect integration of financial markets or goods markets? In Hafer, R., editor, *How open is the US economy?* Lexington Books.

Frenkel, J. (1978). Purchasing power parity: Doctrinal perspective and evidence from the 1920s. *Journal of International Economics*, 8(2):169–191.

Froot, K. and Rogoff, K. (1984). Perspectives on PPP and long-run real exchange rates. NBER Working Paper 4952.

Genberg, H. (1975). *World inflation and the small open economy*, volume 17. Swedish Industrial Publications, Stockholm.

Giovannetti, G. (1987). Testing PPP as a long run equilibrium condition. *Giornale degli Economisti*, XLVI:491–508.

Granger, C. W. J. and Teräsvirta, T. (1993). *Modelling nonlinear economic relationships*. Oxford University Press, Oxford [etc.].

Grossmann, A. and McMillan, D. G. (2010). Forecasting exchange rates: non-linear adjustment and time-varying equilibrium. *Journal of International Financial Markets, Institutions and Money*, 20(4):436–450.

Grossmann, A., Simpson, M. W., and Ozuna, T. (2014). Investigating the PPP hypothesis using constructed U.S. dollar equilibrium exchange rate misalignments over the post-Bretton Woods period. *Journal of Economics and Finance*, 38(2):235–268.

Grunwald, J. and Salazar-Carrillo, J. (1972). Economic integration, rates of exchange, and value comparisons in Latin America.

Hakkio, C. S. (1984). A reexamination of purchasing power parity. *Journal of International Economics*, 17:265–277.

——— (1992). Is purchasing power parity a useful guide to the dollar? *Economic Review*, 77(3):37–51.

Hakkio, C. S. and Rush, M. (1989). Market efficiency and cointegration: an application to the Sterling and Deutschemark exchange markets. *Journal of International Money and Finance*, 8(1):75–88.

Halikias, I., Swagel, P., and Allan, W. (1999). Greece: Selected issues. IMF Staff Country Report 99/138, IMF.

Heckscher, E. F. (1916). Växelkursens grundval vid pappersmyntfot. *Ekonomisk Tidskrift*, 18(10):309–312.

Hsieh, D. A. (1982). The determination of the real exchange rate. *Journal of International Economics*, 12(3):355–362.

Huizinga, J. (1987). An empirical investigation of the long-run behavior of real exchange rates. volume 27, pages 149–214. Carnegie-Rochester Conference Series on Public Policy.

Isard, P. (1977). How far can we push the "law of one price"? *The American Economic Review*, 67(5):942–948.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59:1551–1581.

Kapetanios, G., Shin, Y., and Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics*, 112(2):359.

Kim, Y. (1990). Purchasing power parity in the long run: A cointegration approach. *Journal of Money, Credit and Banking*, 22(4):491–503.

Krugman, P. (1986). Pricing to market when the exchange rate changes. *NBER Working Paper Series*, w1926.

——— (1998). Space: The final frontier. *The Journal of Economic Perspectives*, 12(2):161–174.

Liew, V. K. (2004). Nonlinear adjustment of ASEAN-5 real exchange rates: Symmetrical or asymmetrical? *Economics Bulletin*, 6(1):1–19.

Lundbergh, S., Teräsvirta, T., and van Dijk, D. (2003). Time-varying smooth transition autoregressive models. *Journal of Business and Economic Statistics*, 21(1):104–121.

Luukkonen, R., Saikkonen, P., and Teräsvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika*, 75(3):491–499.

Mark, N. C. (1990). Real and nominal exchange rates in the long run: An empirical investigation. *Journal of International Economics*, 28(1):115–136.

Marston, R. C. (1989). Real exchange rates and productivity growth in the United States and Japan. *NBER Working Paper Series*, w1922.

Michael, P., Nobay, A. R., and Peel, D. A. (1994). Purchasing power parity yet again: Evidence from spatially separated commodity markets. *Journal of International Money and Finance*, 13(6):637–657.

——— (1997). Transactions costs and nonlinear adjustment in real exchange rates: An empirical investigation. *The Journal of Political Economy*, 105(4):862–879.

Milesi-Ferretti, G. M. and Micossi, S. (1994). Real exchange rates and the prices of nontradable goods. *International Monetary Fund, IMF Working Papers: 94/19, 1994*.

Miljkovic, D. (1999). The law of one price in international trade: A critical review. *Review of Agricultural Economics*, 21(1):126–139.

Obstfeld, M. and Rogoff, K. (2000). The six major puzzles in international macroeconomics: Is there a common cause? *NBER Macroeconomics Annual*, 15(1):339–390.

Obstfeld, M. and Taylor, A. M. (1997). Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited. *Journal of the Japanese and International Economies*, 11(4):441–479.

Officer, L. H. (1976). The productivity bias in purchasing power parity: An econometric investigation. *Staff Papers*, 23(3):545–579.

Pakko, M. R. and Pollard, P. S. (2003). Burgernomics: A Big Mac™ guide to purchasing power parity.

Rogoff, K. (1996). The purchasing power parity puzzle. *Journal of Economic Literature*, 34:647–668.

Roll, R. (1979). Violations of purchasing power parity and their implications for efficient international commodity markets. In Sarnat, M. and Szego, G., editors, *International Finance and Trade*. Ballinger.

Rother, C. (2000). The impact of productivity differentials on inflation and the real exchange rate: An estimation of the Balassa–Samuelson effect in Slovenia. IMF Staff Country Report 00/56, IMF. pp.26-38.

Samuelson, P. A. (1954). The transfer problem and transport costs: Analysis of effects of trade impediments. *The Economic Journal*, 64(254):264–289.

——— (1964). Theoretical notes on trade problems. *The Review of Economics and Statistics*, 46(2):145–154.

Strauss, J. (1995). Real exchange rates, PPP and the relative price of nontraded goods. *Southern Economic Journal*, 61(4):991–1005.

Taylor, A. M. (2001). Potential pitfalls for the purchasing-power-parity puzzle?: Sampling and specification biases in mean-reversion tests of the law of one price. *Econometrica*, 69(2):473–498.

Taylor, M. P. (1988). An empirical examination of long-run purchasing power parity using cointegration techniques. *Applied Economics*, 20(10):1369–1381.

Taylor, M. P. and McMahon, P. C. (1988). Long-run purchasing power parity in the 1920s. *European Economic Review*, 32(1):179–197.

Taylor, M. P., Peel, D. A., and Sarno, L. (2001). Nonlinear mean-reversion in real exchange rates: Toward a solution to the purchasing power parity puzzles. *International Economic Review*, 42(4):1015–1042.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425):208–218.

——— (1998). Modelling economic relationships with smooth transition regressions.

Teräsvirta, T. and Anderson, H. M. (1992). Characterizing nonlinearities in business cycles using smooth transition autoregressive models. *Journal of Applied Econometrics*, 7(S1):S119–S136.

The Economist (2014). The Big Mac index.

Tica, J. and Druzic, I. (2006). The Harrod-Balassa-Samuelson effect: A survey of empirical evidence. FEB–Working Paper Series 06-7, University of Zagreb.

Tong, H. (1990). *Non-linear time series: a dynamical system approach*, volume 6. Clarendon Press, Oxford [etc.].

Van Dijk, D. and Franses, P. H. (2006). A simple test for PPP among traded goods. *Applied Financial Economics*, 16(1/2):19–27.

Wei, S. and Parsley, D. C. (1995). Purchasing power disparity during the floating rate period: Exchange rate volatility, trade barriers and other culprits. *NBER Working Paper Series*, w5032.

Xu, Z. (2003). Purchasing power parity, price indices, and exchange rate forecasts. *Journal of International Money and Finance*, 22(1):105–130.

6 Appendix

Table A.1: Preliminary test: ADF unit root test

Series	ADF t-Statistic	Regression specification	Order of Integration
Canada			
Spot EXR	-1.35	Intercept, 1 lag	1
CPI	-7.74***	Intercept & Trend, no lags	0
TPI	-2.42	Intercept & Trend, 2 lags	1
Japan			
Spot EXR	-2.09	Intercept & Trend, 1 lag	1
CPI	-1.51	Intercept & Trend, 12 lags	1
TPI	-2.57	Intercept & Trend, 3 lags	1
Netherlands			
Spot EXR	-3.33*	Intercept & Trend, no lags	1
CPI	-3.99***	Intercept & Trend, 12 lags	0
TPI	-1.53	Intercept, 1 lag	1
Singapore			
Spot EXR	-1.89	Intercept & Trend, 1 lag	1
CPI	-2.29	Intercept & Trend, 12 lags	1
TPI	-1.98	Intercept, 1 lag	1
United States			
CPI	-3.57**	Intercept & Trend, 2 lags	0
TPI	-2.49	Intercept & Trend, 1 lag	1

Results for the Augmented Dickey Fuller test on the spot exchange rate (spot EXR) and price index (CPI and TPI) series. *,** and *** denote significance at the 10, 5 and 1 percent level respectively. The specification shows the added components in the ADF regression, the number of lags is determined with Schwarz Information Criterion. Order of integration displays whether the series is perceived as stationary (0) or nonstationary (1).

Table A.2: KSS δ supplement

Country	Estimated δ	t-Statistic	Simulated probability
CPI			
Canada	-0.30	-1.74	0.1924
Japan	-0.16	-2.32	0.0884
Netherlands	-0.01	-0.40	0.8320
Singapore	-0.14	-0.79	0.6556
TPI			
Canada	-0.11	-1.55	0.2724
Japan	-0.04	-1.97	0.1294
Netherlands	-0.01	-1.04	0.5282
Singapore	-0.26	-2.15	0.0988

The estimated ESTAR δ coefficients from table 1, from the specification in equation (23) estimated using Ordinary Least Squares. Under $H_0 : \delta = 0$ equation (23) is a random walk, for which the student t-distribution is not valid. This invalidates the probabilities that are given by the econometric software. The solution is to construct a simulated distribution for the t-Statistic when the actual underlying process is a random walk. The first step is to simulate a random walk and estimate KSS specifications for that data. Each random walk simulation consisted of 508 observations of which the first 100 were dropped so as to match this report's observation count of 408. Thus, over these remaining 408 random walk observations an KSS equation was estimated and the corresponding t-Statistic of δ was stored. This process was repeated for 5000 simulations. Afterwards, the actual t-Statistics from table 1 were compared with these 5000 simulated t-Statistics. The percentage of simulated t-Statistics that lay above the (absolute) value of the actual t-Statistic was then used as the simulated probability, displayed in the far right column.

Table A.3: ESTAR θ supplement

Country	Estimated θ	t-Statistic	Simulated probability
CPI			
Canada	3.19	1.60	0.1484
Japan	1.71	1.96	0.0436
Netherlands	0.06	0.39	0.8402
Singapore	1.45	0.76	0.6456
TPI			
Canada	1.12	1.40	0.2402
Japan	0.38	1.79	0.0788
Netherlands	0.10	0.98	0.5150
Singapore	3.08	1.90	0.0532

The estimated ESTAR θ coefficients from table 2, from the specification in equation (19) estimated using Nonlinear Least Squares. Under $H_0 : \theta = 0$ equation (19) is a random walk, for which the student t-distribution is not valid. This invalidates the probabilities that are given by the econometric software. The solution is to construct a simulated distribution for the t-Statistic when the actual underlying process is a random walk. The first step is to simulate a random walk and estimate ESTAR specifications for that data. Each random walk simulation consisted of 508 observations of which the first 100 were dropped so as to match this report's observation count of 408. Thus, over these remaining 408 random walk observations an ESTAR equation was estimated and the corresponding t-Statistic of θ was stored. This process was repeated for 5000 simulations. Afterwards, the actual t-Statistics from table 2 were compared with these 5000 simulated t-Statistics. The percentage of simulated t-Statistics that lay above the (absolute) value of the actual t-Statistic was then used as the simulated probability, displayed in the far right column.

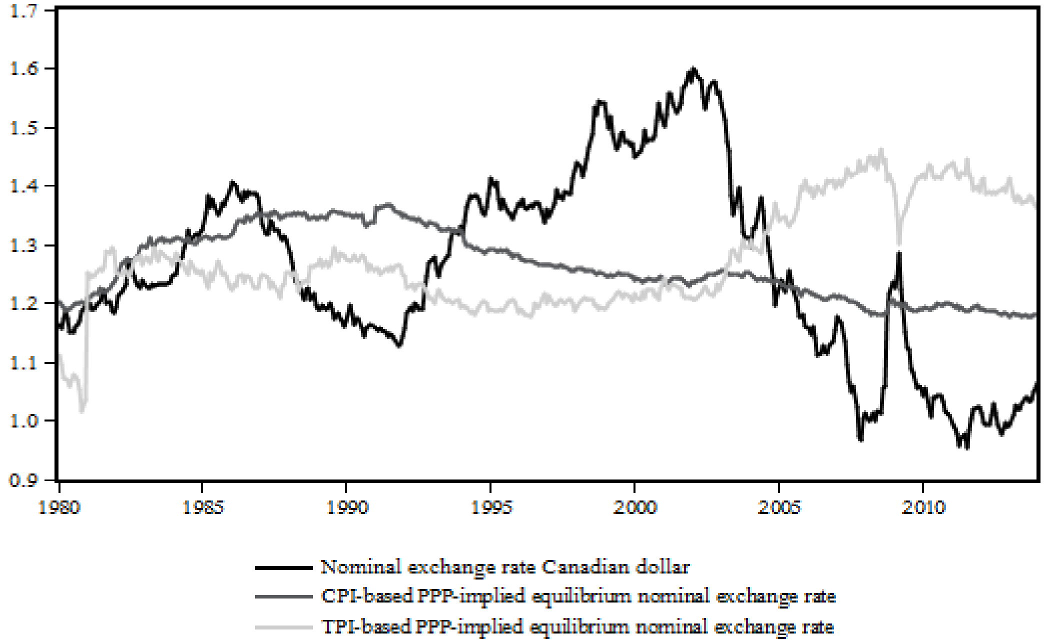

Figure A.1: Canadian dollar/US dollar exchange rate and the associated CPI and TPI PPP-implied equilibria.

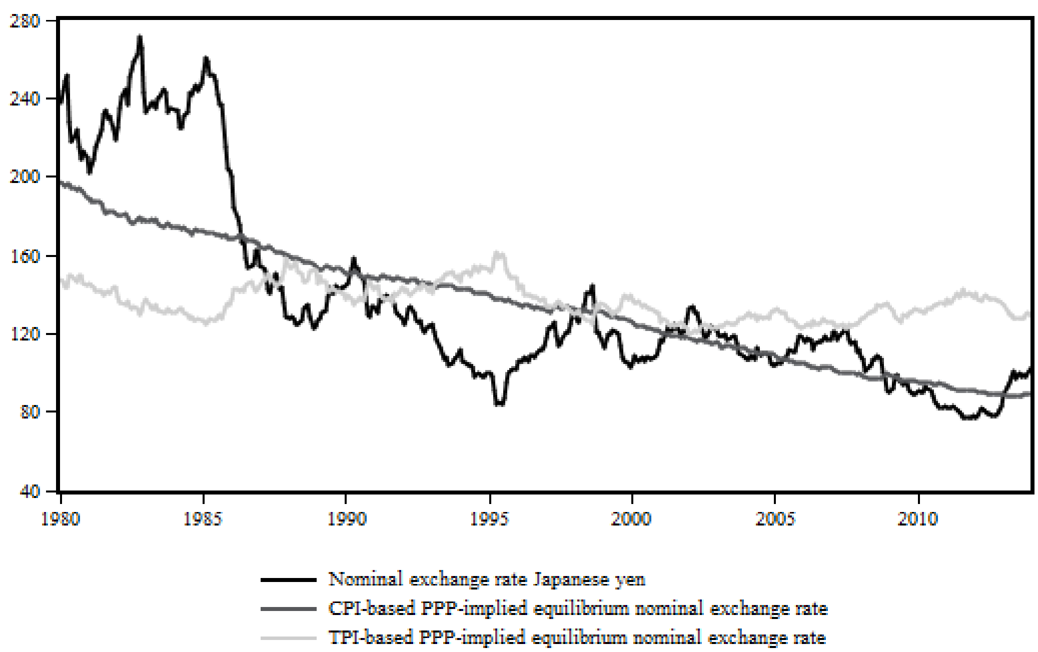

Figure A.2: Japanese yen/US dollar exchange rate and the associated CPI and TPI PPP-implied equilibria.

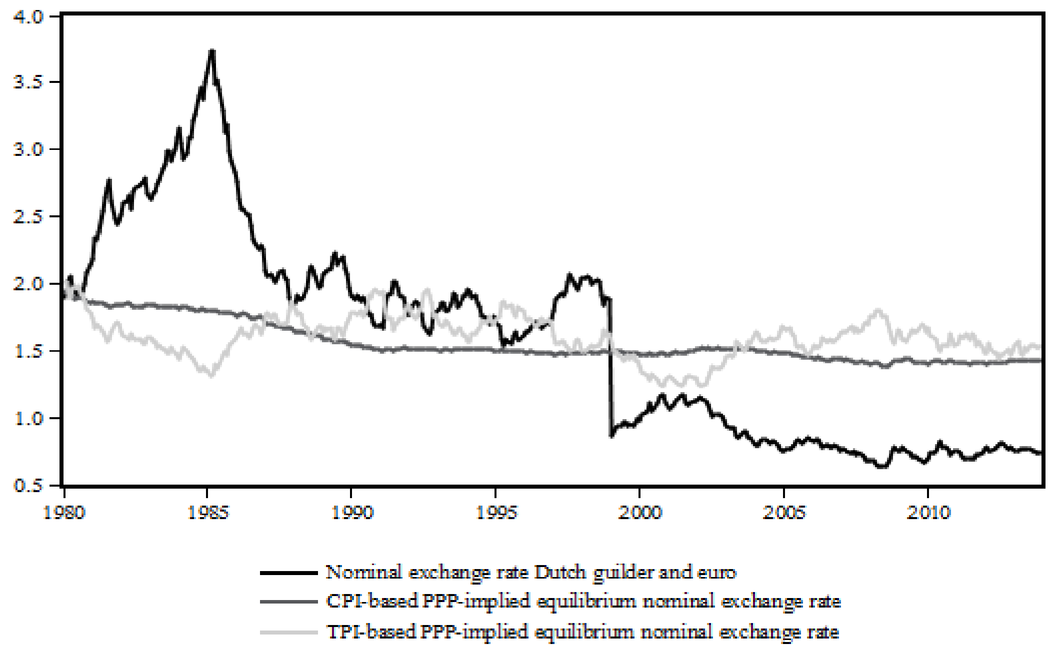

Figure A.3: Dutch guilder(euro)/US dollar exchange rate and the associated CPI and TPI PPP-implied equilibria.

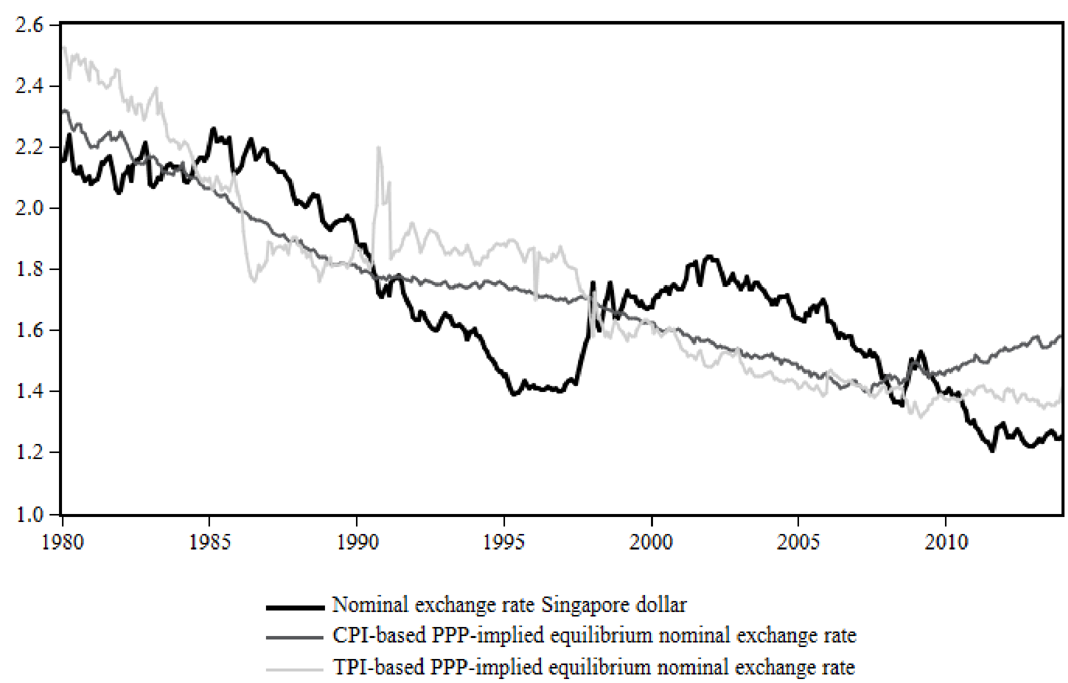

Figure A.4: Singapore dollar/US dollar exchange rate and the associated CPI and TPI PPP-implied equilibria.

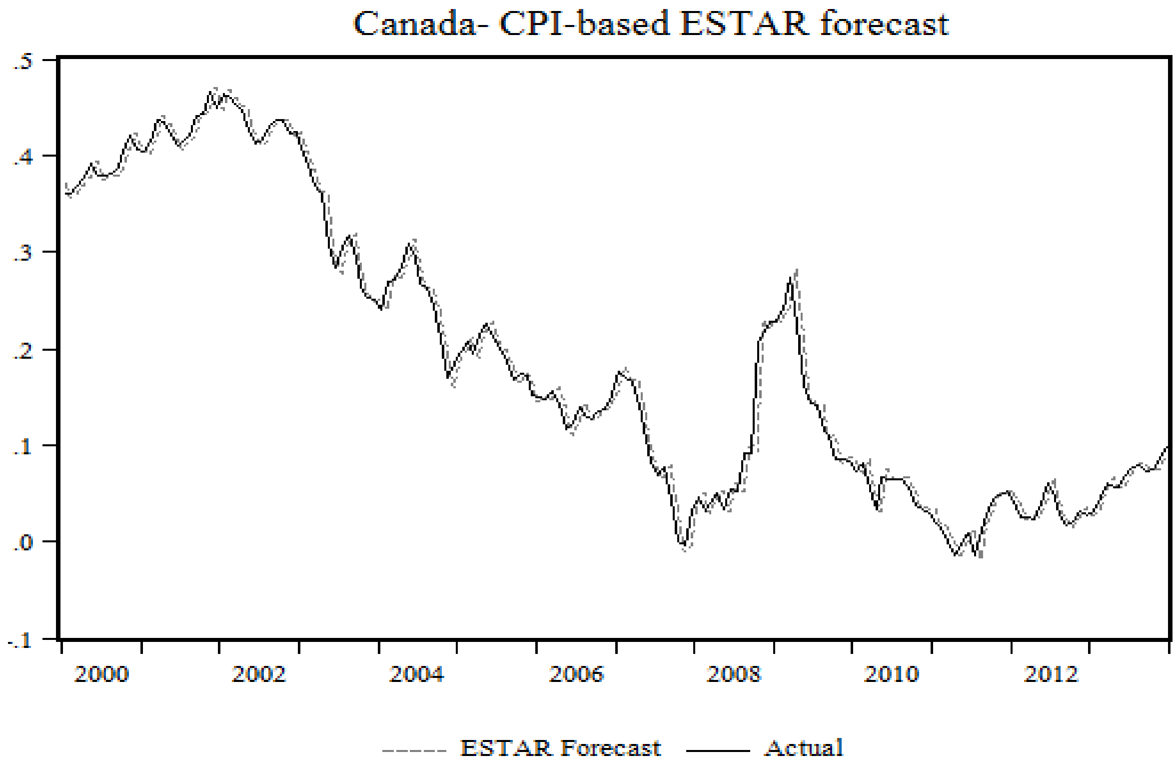

Figure A.5: The actual Canadian dollar real exchange rate versus the recursive CPI-based ESTAR forecast.

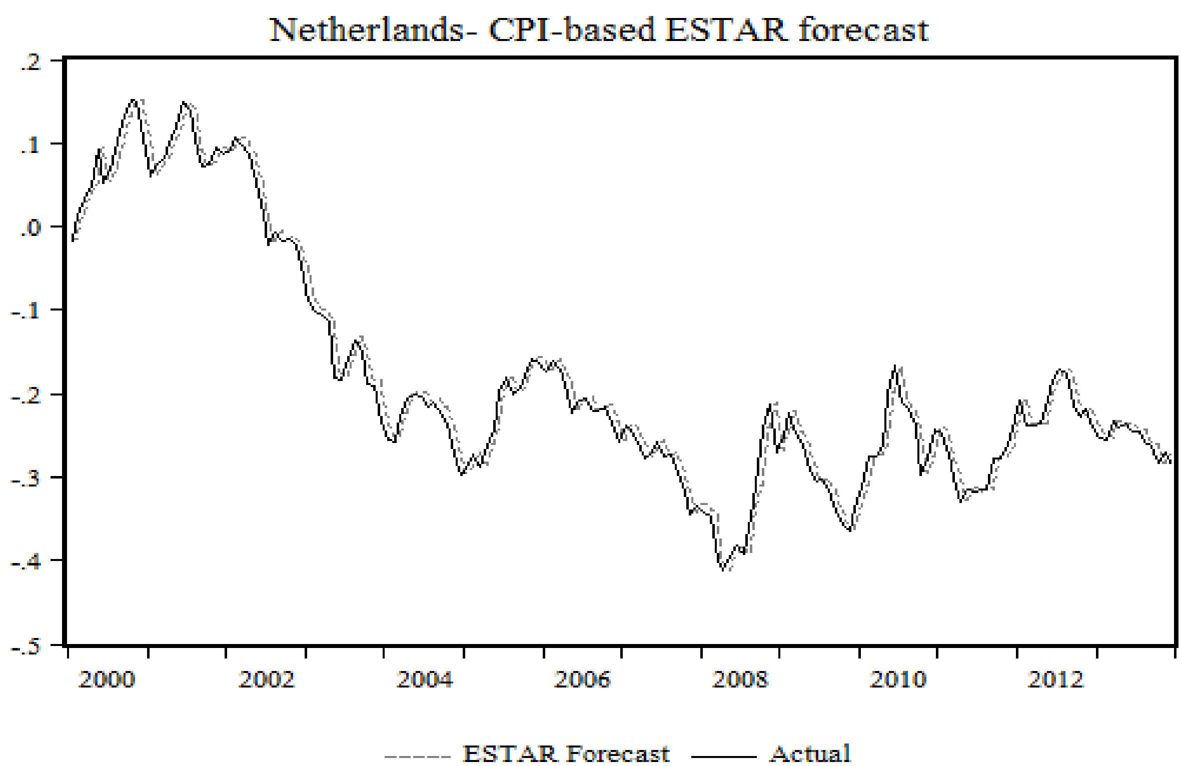

Figure A.6: The actual euro real exchange rate versus the recursive CPI-based ESTAR forecast.

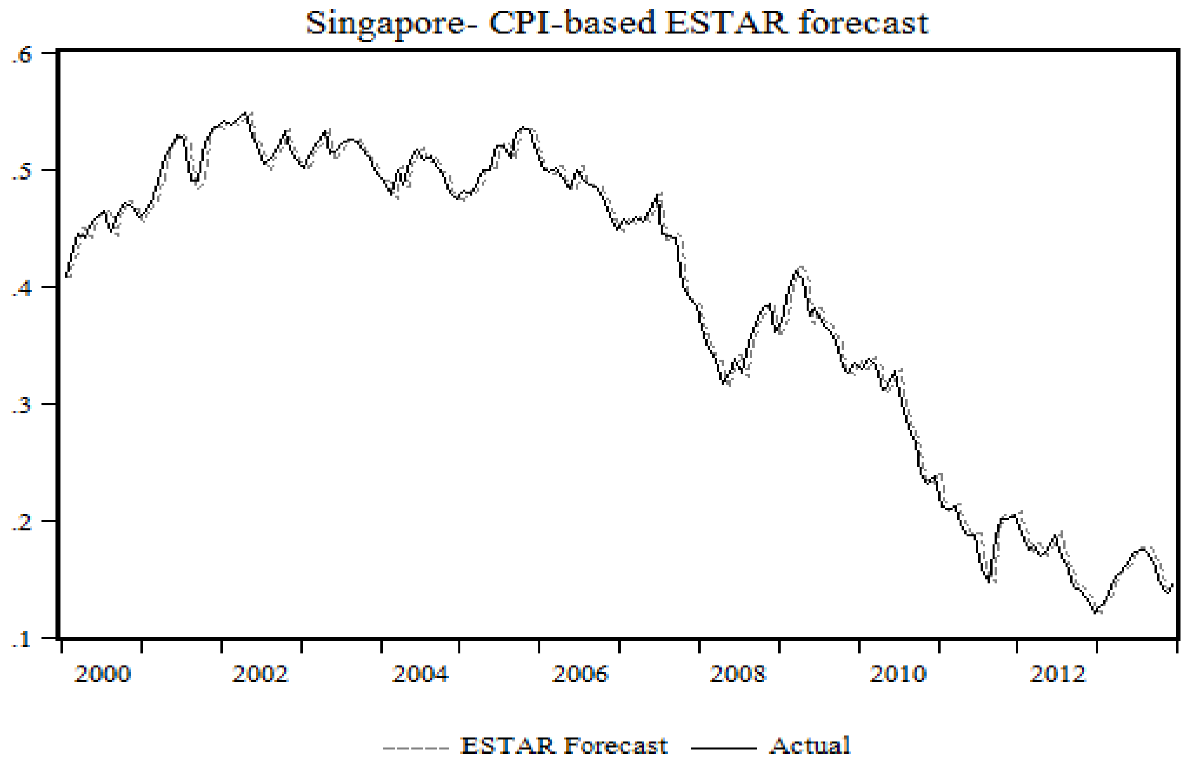

Figure A.7: The actual Singapore dollar real exchange rate versus the recursive CPI-based ESTAR forecast.

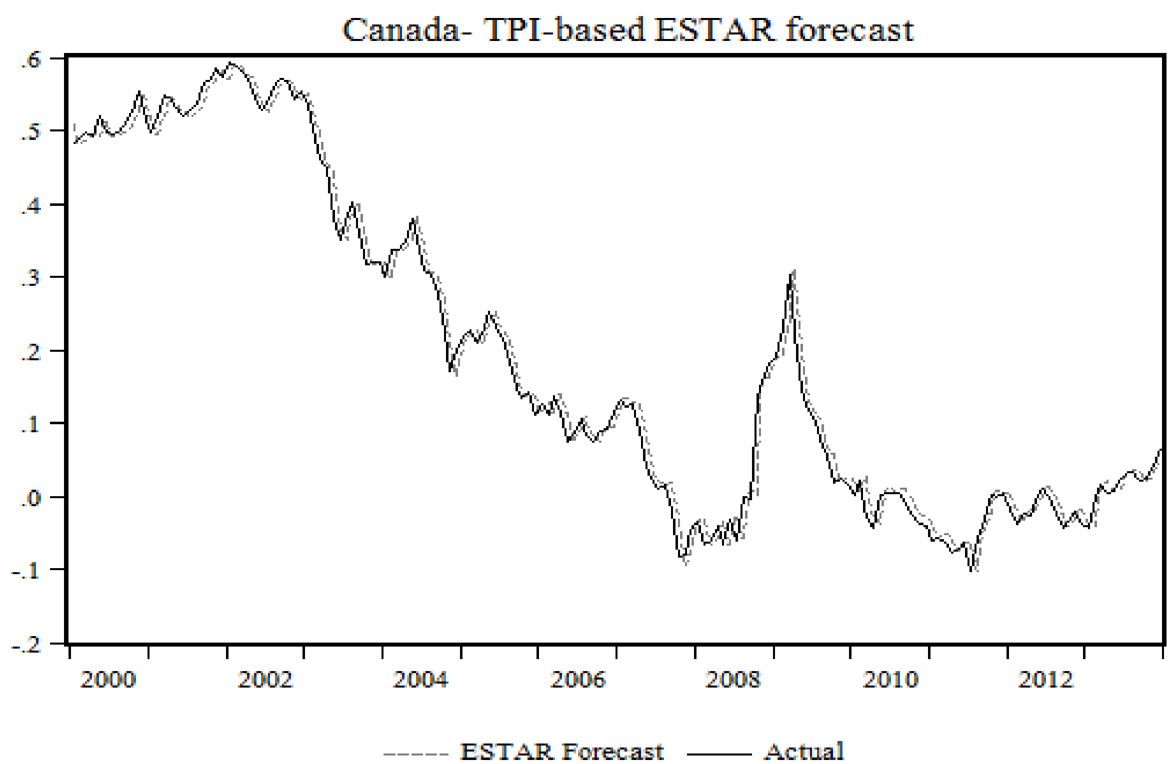

Figure A.8: The actual Canadian dollar real exchange rate versus the recursive TPI-based ESTAR forecast.

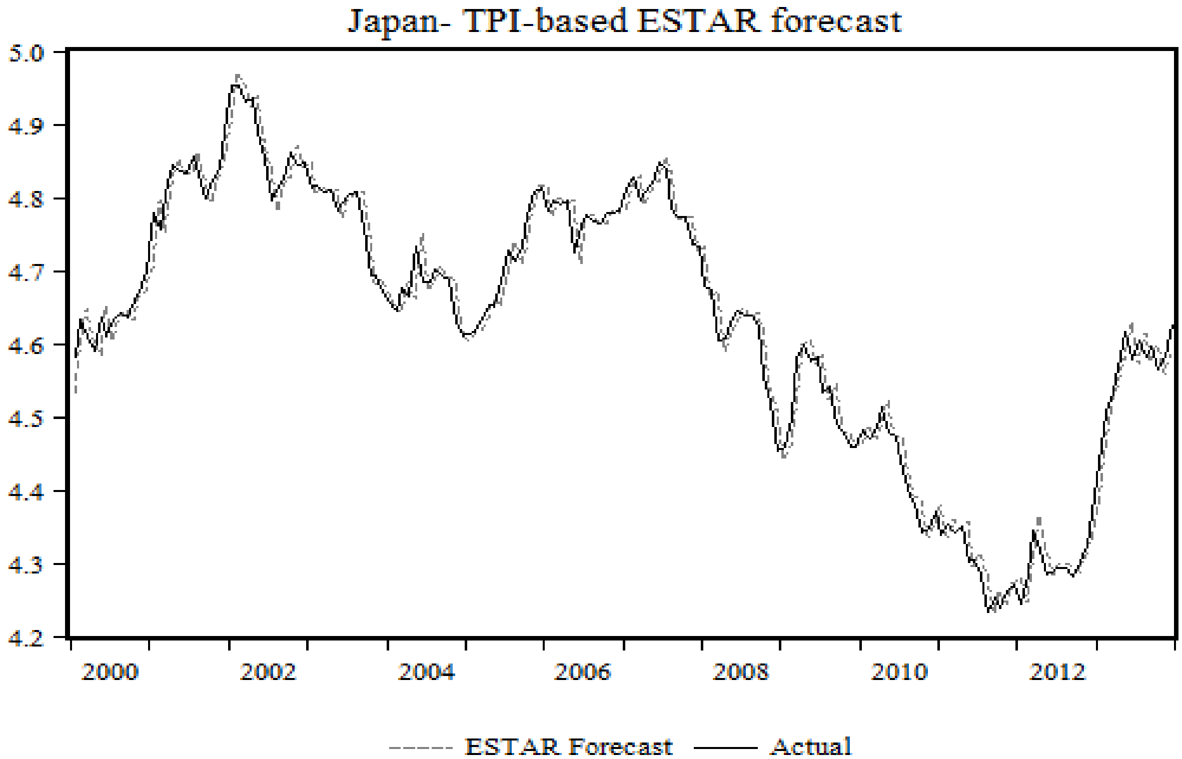

Figure A.9: The actual Japanese yen real exchange rate versus the recursive TPI-based ESTAR forecast.

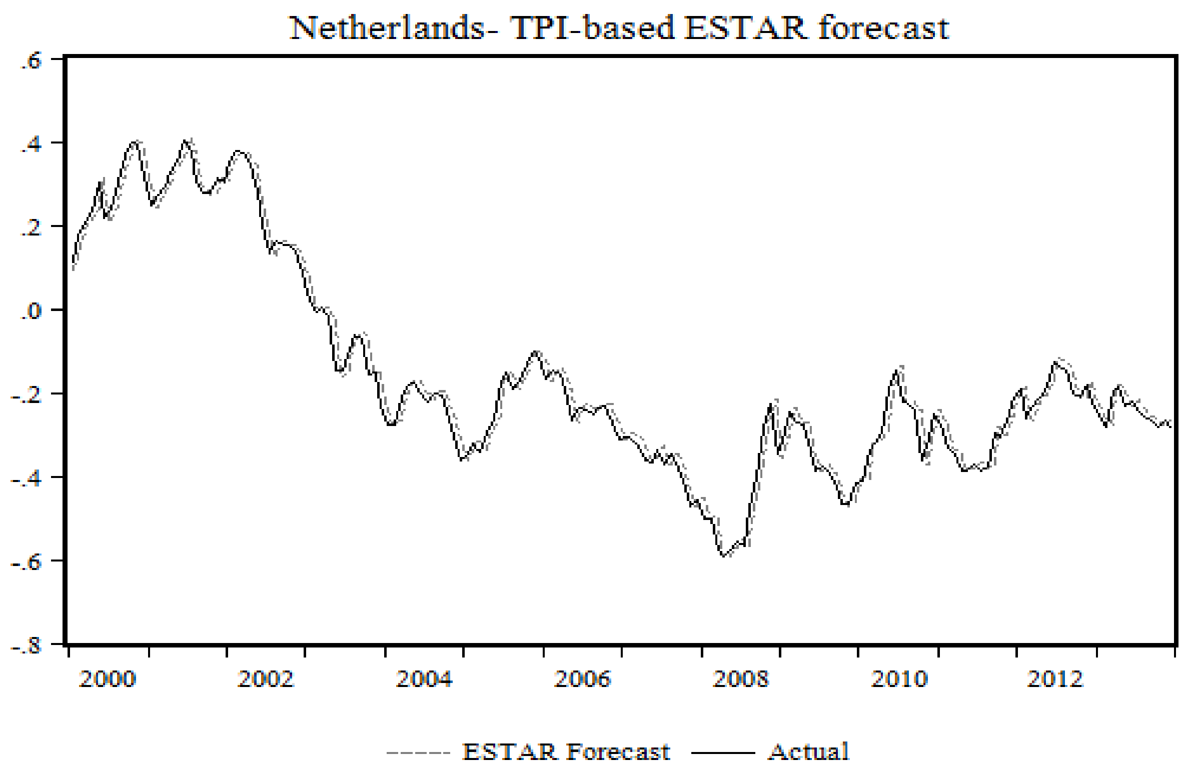

Figure A.10: The actual euro real exchange rate versus the recursive TPI-based ESTAR forecast.

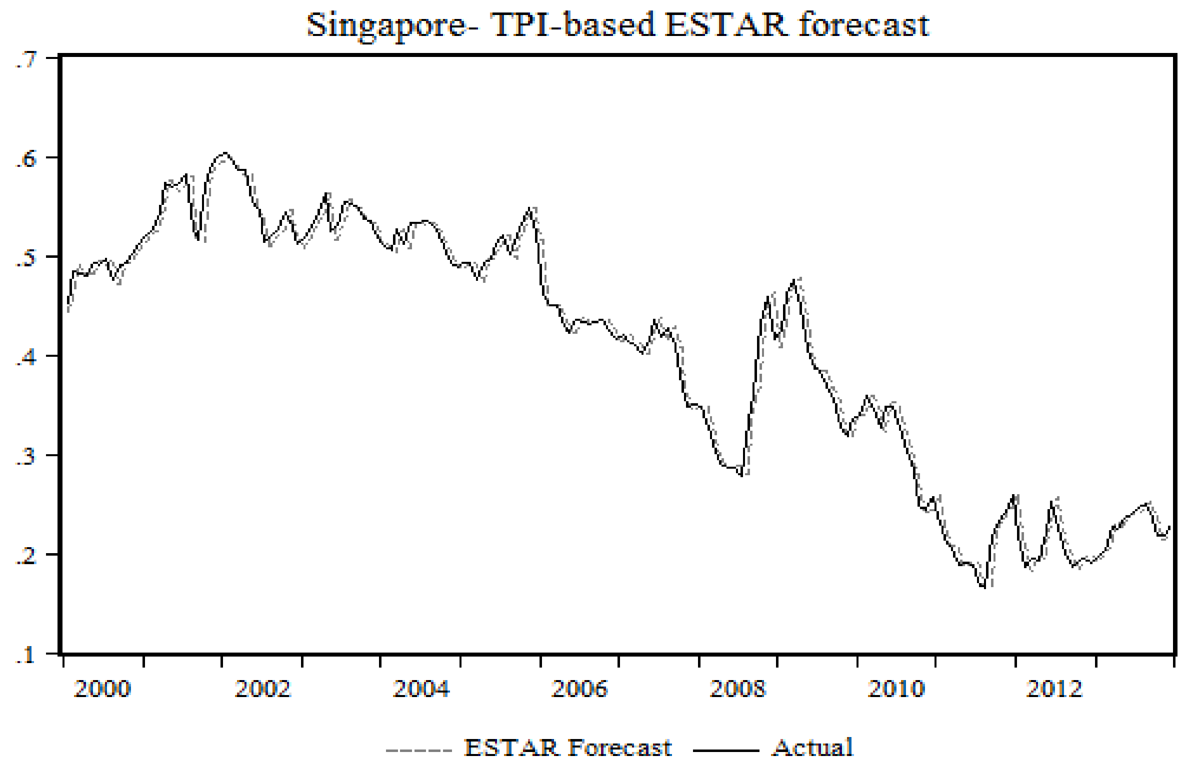

Figure A.11: The actual Singapore dollar real exchange rate versus the recursive TPI-based ESTAR forecast.