THE RIMINI CENTRE FOR ECONOMIC ANALYSIS

WP 15-43

Richard T. Baillie
Michigan State University, USA
Queen Mary University of London, UK
The Rimini Centre for Economic Analysis, Italy

Kun Ho Kim
Hanyang University, Republic of Korea

LOCAL DEVIATIONS FROM UNCOVERED INTEREST PARITY: THE ROLE OF MACROECONOMIC FUNDAMENTALS

The Rimini Centre for Economic Analysis
Legal address: Via Angherà, 22 – Head office: Via Patara, 3 – 47921 Rimini (RN) – Italy
www.rcfea.org – secretary@rcfea.org

Local Deviations from Uncovered Interest Parity: The Role of Macroeconomic Fundamentals

Richard T. Baillie

Department of Economics, Michigan State University, USA

School of Economics and Finance, Queen Mary, University of London, UK

Rimini Center for Economic Analysis, Italy

and

Kun Ho Kim

Department of Economics and Finance, Hanyang University, Seoul, Korea

April 29, 2015

Abstract

This paper uses recently developed kernel smoothing regression procedures and uniform confidence bounds to investigate the forward premium anomaly. These new statistical methods estimate the local time varying slope coefficient of the regression of spot returns on the lagged interest rate differential. The uniform confidence bands indicate the extent of the rejections of uncovered interest parity and find remarkable variation in both regimes when the anomaly occurs, and also the magnitude of the slope coefficient estimate. Of particular interest is the fact that the time varying slope parameter can be substantially explained by fundamentals such as monetary growth rates, and also the volatility of US money growth, which is associated with risk premium in many theoretical mdoels. Hence, the apparent deviations from uncovered interest parity have explanations consistent with monetary models and associated risk premium models.

Key words: Forward premium anomaly, Local Deviation from Uncovered Interest Parity, Time-varying parameters, Kernel smoothing, Uniform inference, Macroeconomic fundamentals, Model averaging.

JEL Classifications: C12; C14; C22; F31; F41; G15

1 Introduction

One of the long standing issues in international finance has been the apparent failure of the theory of Uncovered Interest Rate Parity (UIP). The classic method for testing UIP is to estimate the slope coefficient in a regression of spot returns on the lagged forward premium, or equivalently, the lagged interest rate differential. While the slope coefficient should be unity under UIP, most studies have found statistically significant rejections of the UIP hypothesis, with the slope coefficient estimate invariably being quite large and negative. This has become known as the *forward premium anomaly.* Hence most research has been directed at understanding the reasons for the apparent rejection of UIP and to try to account for it in terms of (i) time dependent risk premium, (ii) irrational agents and segmented markets, (iii) peso problems, or (iv) econometric issues with the testing of UIP. The dominant approach has been to explain the phenomenon by modeling a time dependent risk premium. Overall, this approach has not been particularly successful empirically.

The theory of UIP under rational expectations and a constant risk premium implies that

$$E_t(\Delta s_{t+1}) = (f_t - s_t) = (i_t - i_t^*) \tag{1}$$

is always an approximation which neglects the Jensen inequality terms, and possible time dependent risk premium. It has become standard to test the theory from the regression equation

$$\Delta s_{t+1} = \alpha + \beta(f_t - s_t) + u_{t+1}, \tag{2}$$

where the theory of UIP implies $\alpha = 0$, $\beta = 1$ and u_{t+1} being serially uncorrelated[1].

[1]Some studies such as Hansen and Hodrick (1980), Hakkio (1981) and Baillie, Lippens and Mcmahon (1983) tested the theory with overlapping data where the maturity time of the forward contract exceeds the sampling interval of the data. These studies still find rejection of UIP.

However, an increasing number of studies have come to recognize the fact that departures from UIP are more pronounced in some periods than others. The usual way of representing the potential variation in the slope coefficient is by *rolling regressions*, as in Baillie and Bollerslev (2000), Lothian and Wu (2011), etc. Other studies by Wolf (1987) have used Kalman filtering with the β following a random walk or stationary autoregression; while Bansal (1997) and Bansal and Dahlquist (2000) have allowed β to have two states depending on the sign of the interest rate differental; and Baillie and Kilic (2006) use a logistic smooth transiton regression to allow the β parameter to move slowly between the two states which correspond to either UIP holding[2], or alternatively a state with the forward premium anomaly being apparent. These parametric specifications for the time series behavior of the slope coefficient over time are necessarily heavily dependent on the parametric specification of the time series process for β_t.

While simple to apply in practice, the rolling regression technique is, however, highly arbitrary in the sense that the number of observations used in the window is very subjective. That is, there is no dependable criterion that one can use in choosing the right window size. The method also tends to produce quite wide confidence intervals from OLS regressions but does not allow any clear mehod for conducting statistical inference between different regressions.

---

[2]It should be noted that a more general representation of UIP is to begin with a standard discrete time, consumption based asset pricing model where the real returns of the representative investor are $E_t\left(\frac{S_{t+1}-F_t}{P_{t+1}}\right)\left(\frac{U'(C_{t+1})}{U'(C_t)}\right)=0$, where S and F are the spot exchange rate and forward rate in levels, P is domestic price level and C is domestic consumption, and $U'(C_t)$ is the marginal utility of consumption in period t. Then,

$$E_t(\Delta s_{t+1}) = (f_t - s_t) - \left(\frac{1}{2}\right) Var_t(\Delta s_{t+1}) + Cov_t\left(\Delta s_{t+1} p_{t+1}\right) + \rho_t, \tag{3}$$

where ρ_t is the natural logarithm of the intertemporal marginal rate of substitution and is generally called the "risk premium". The above theory dates back at least to Hansen and Hodrick. (1983)

One major novelty in this paper is to introduce the concept of Local Deviation from Uncovered Interest Parity ($LDUIP$), which is the specific amount that the parity condition is violated at each time point and is based on non-parametric and local smoothing techniques developed for the *local-linear regression* introduced by Stone (1977) and by Cleveland (1979). These techniques avoid the problems with rolling regressions and produce kernel smoothed regressions. They also allow *statistical inference* to be conducted on the parameters. The method assumes that the regression parameters are *smoothly varying* functions of time, and circumvents possible abrupt and sudden changes in the parameters. The method also enables the construction of *uniform confidence bands* (UCB) of the slope coefficient from its local-linear regression estimate. Hence the slope coefficient of the forward premium regression can be tested for any parametric specification of the unknown function. The generated $LDUIP$ process and its associated UCB indicate the extent and significance of possible violations of UIP at any point of time.

A further interesting issue centers on the reasons for the changes in the relatively smooth pattern of the $LDUIP$ and to what extent they can be predicted by macroeconomic fundamentals and, or variables associated with time dependent risk premium. Some evidence is presented in section 4 of the paper that indicates a substantial role for lagged macroeconomic fundamentals and variables associated time risk premium, to have predictive power in explaining the movements of the time varying parameter in the forward premium regression. Given that the $LDUIP$ are in some sense "model free", deterministic estimates of the slope parameter in the UIP regression, interest focuses on the reasons for the time variation in these coefficients. Various fundamental based explanations and also models based on some models developed to explain time-dependent risk premium are used in a second step analysis, where the generated deterministic β_t are regressed on five different risk premium models, which typically contain estimated second moments, or conditional variances and covariances of some variables associated with previously developed economic models of risk premium. The validity and relative strength of each model is then assesses through a classical frequentist based model averaging based procedure. This method indicates the most likely reason for the breakdown of UIP over the whole sample and also for certain sub periods such as the financial crisis.

The organization of the paper is the following: Section 2 introduces the model framework, and the forward premium regression with smoothly varying coefficients is explained. Section 3 then discusses the kernel smoothing regression and the construction of uniform confidence bands (UCB) for inference. Section 4 presents the empirical results including the estimates of the time varying, estimated slope coefficients. The $UCBs$ determine the precise time and extent of the violation of UIP for each currency over the sample period. This section also includes evidence from regression tests and $VARs$ on the role of some fundamentals and financial variables that appear related to changes in the slope parameter of the forward premium regression. Section 5 concludes the paper and discusses related future research. The technical assumptions on the model and the details of the steps to construct the UCB are relegated to an appendix.

2 Model framework

It is first worth noting that apart from the strong empirical evidence, there are also economic reasons to allow the conventional regression in equation (2) to have time varying parameters. For example, the model

$$\Delta s_{t+1} = \alpha_t + \beta_t (f_t - s_t) + u_{t+1}, \tag{4}$$

can be justified from the approach of Chang (2013), where there is cross country speculation in stocks and bonds. Then with financial traders with negative exponential utility, the model implies a β_t that is the population equivalent of a regression of spot returns on lagged equity returns differentials. Extension of the model to include risk, or expectational error leads to a β_t that is the population equivalent of a regression of spot returns on a linear combination of variables associated with risk or expectational errors. Another motivation for time varying β_t could involve adaptation of the Taylor rule used in the exchange rate model of Engel and West (2005).

The approach taken on this paper is to be agnostic about the possible reasons for time variation in the slope coefficient and to essentially model it without any strong restrictions. Hence we estimate a model of the form

$$s_{\tau+1} - s_\tau = \beta_0 + \beta_1(f_\tau - s_\tau) + \epsilon_{\tau+1}, \qquad \tau = 1, \cdots, T-1. \tag{5}$$

where s_τ is the log of the monthly spot exchange rate at time τ, quoted as the foreign price of domestic currency, while f_τ is the log of the corresponding 30 day forward rate, and finally $\epsilon_{\tau+1}$ is the stationary and serially uncorrelated disturbance term. In the above and throughout the paper, the index τ is reserved to indicate *discrete* time. Since the underlying unknown parameters β_0 and β_1 are specified as being deterministic and continuous functions of time, the traditional time index t is reserved for the continuous-time processes in $[0, 1]$. Hence the following specification with *time-varying* parameters is then introduced to replace the traditional forward premium regression,

$$s_{\tau+1} - s_\tau = \beta_0\left(\frac{\tau}{T}\right) + \beta_1\left(\frac{\tau}{T}\right)(f_\tau - s_\tau) + \epsilon_{\tau+1}, \qquad \tau = 1, \cdots, T-1. \tag{6}$$

where $\beta_0(\cdot)$ and $\beta_1(\cdot)$ are non-parametric functions of time which allows for potential time variation in the model parameters. Another advantage of this framework is that it enables simultaneous inference to be implemented for the unknown parameter functions.

Moreover, it is assumed that the regressors in (6) are locally stationary variables (Zhou and Wu, 2010; Kim, 2014a; Kim, 2014b). In a relatively short time span, they are approximately stationary. However, as the time horizon increases the variables show the characteristics of non-stationary processes, such as time-varying moments. One of the appealing features for this flexible class of non-stationarity is that its framework can encompass a wide range of popular linear or non-linear time series processes, such as various stationary processes or auto-regressive processes with time varying parameters. Further details of the assumptions for (6) are given in Appendix 1.

One of the main features of the analysis in this paper is that, apart from modeling the time variation in the $\beta_1(\cdot)$, it also provides quite precise UCB that clearly indicate statistically significant departures from UIP. The precise reason for departures from UIP and their relationship to macroeconomic fundamentals and time-dependent risk premium are investigated in section 4. Suffice to say, that these macro and financial variables are found to play an important role in the time-varying nature of the slope coefficient estimate in the forward premium regression.

3 Methodology

The forward premium regression (6) with the smoothly time-varying coefficients, are estimated by using kernel-smoothing techniques. This method allows one to perform estimation and inference of the true underlying process without imposing any arbitrary parametric assumptions on it. Thus, it minimizes the possibility of mis-specification. In addition, the method is computationally tractable.

3.1 Local-linear regression

For simplicity, let $\Delta s_{\tau+1} := s_{\tau+1} - s_\tau$, $\mathbf{x}_\tau^{/} := [1 \;\; f_\tau - s_\tau]$ and $\boldsymbol{\beta}(t)^{/} := [\beta_0(t) \;\; \beta_1(t)]$, where $0 \leq t \leq 1$. Among various kernel smoothing techniques, the local-linear regression method of Cleveland (1979) stands out due to its simple form, ease of computation and analytical tractability. In contrast to other popular kernel smoothing methods, such as the Nadaraya-Watson estimator, the local-linear regression estimator does suppress the boundary problem and achieve nearly optimal statistical efficiency. The local-linear regression estimates of the parameters in (6) are obtained by the following optimization:

$$(\hat{\boldsymbol{\beta}}(t), \hat{\boldsymbol{\beta}}'(t)) := \underset{(\eta_0,\eta_1)}{\operatorname{argmin}} \sum_{\tau=1}^{T-1} \left[\Delta s_{\tau+1} - \mathbf{x}_\tau^{/}\eta_0 - \mathbf{x}_\tau^{/}\eta_1 \left(t - (\tau/T)\right)\right]^2 K\left(\frac{t-(\tau/T)}{b}\right) \tag{7}$$

where $\hat{\boldsymbol{\beta}}(t)^{/} = [\hat{\beta}_0(t) \;\; \hat{\beta}_1(t)]$ are the estimates of the model coefficients, $\beta_0(t)$ and $\beta_1(t)$, and $\hat{\boldsymbol{\beta}}'(t)^{/} = [\hat{\beta}_0'(t) \;\; \hat{\beta}_1'(t)]$ are the estimates of their first-order derivatives, $\beta_0'(t)$ and $\beta_1'(t)$. Here $K(\cdot)$ is a kernel function and b is a bandwidth. In this study, we employ the Epanechnikov kernel $K(x) = 3\max(1-x^2, 0)/4$. The bandwidth is chosen by the generalized cross validation (GCV) procedure of Craven and Wahba (1979), and is described in detail in section (i) of Appendix 2. Alternatively, one can instead consider the following *jackknife bias-corrected* estimator (Zhou and Wu, 2010):

$$\tilde{\boldsymbol{\beta}}(t) := 2\hat{\boldsymbol{\beta}}_{b/\sqrt{2}}(t) - \hat{\boldsymbol{\beta}}(t) \tag{8}$$

where $\hat{\boldsymbol{\beta}}_{b/\sqrt{2}}(t)$ is $\hat{\boldsymbol{\beta}}(t)$ with $b/\sqrt{2}$ instead of the original bandwidth b.

A major motivation for using the kernel smoothed regression method in (7) comes from the literature that has investigated the time series properties of the forward premium. In particular, Baillie and Bollerslev (1994, 2000), Maynard and Phillips (2001) and Sakoulis, Zivot and Choi (2010) have found strong evidence that the forward premium, or equivalently the interest rate differential is a long memory, fractionally integrated process. There is also clear evidence that forward premium series are typically highly non-linear, with possible break points. Hence the general class of non-stationary, locally stationary process of Zhou and Wu (2010) and Kim (2014a, 2014b) seems an ideal assumption for dealing with this type of time series [3].

[3]Other possible approaches for dealing with time varying parameters are to be found in the articles by Chen and Tsay (1993), Phillips (2001), Orbe, Ferreira and Rodrigues-Poo (2005)

3.2 Uniform confidence band

One of the important advantages of the kernel smoothed regression in (7) is that the methodology allows the construction of uniform confidence bands (UCB) for the time-varying parameters $\beta_0(\cdot)$ and $\beta_1(\cdot)$ in (6). Given the UCB, one can perform simultaneous inference for the true underlying process, which facilitates the testing of the shape characteristics and subsequent tests for constancy, or linearity, etc.

In order to construct the asymptotic UCB of parameter $\beta_j(t)$, $j = 0, 1$, over $t \in [0, 1]$ with the confidence level $100(1 - \alpha)\%$, $\alpha \in (0, 1)$, one needs to find two functions $\ell_{j,n}(\cdot)$ and $u_{j,n}(\cdot)$ based on data, such that:

$$\lim_{n\to\infty} \mathbb{P}\{\ell_{j,n}(t) \leq \beta_j(t) \leq u_{j,n}(t) \text{ for } \textit{all}\; t \in [0, 1]\} = 1 - \alpha \tag{9}$$

The main purpose of constructing the UCB in (9) is to test whether the parameter $\beta_j(t)$ takes a certain parametric form. That is, using (9), we are able to test the null hypothesis $H_0 : \beta_j(\cdot) = \beta_{j,\theta}(\cdot)$, where $\theta \in \Theta$ and Θ is a parameter space. For example, in order to test $H_0 : \beta_j(t) = \theta_0 + \theta_1 t + \theta_2 t^2$, one can simply check whether $\ell_{j,n}(t) \leq \hat{\theta}_0 + \hat{\theta}_1 t + \hat{\theta}_2 t^2 \leq u_{j,n}(t)$ holds for *all* $t \in [0, 1]$. Here $\hat{\theta}_i$ is the least squares estimate of θ_i, $i = 0, 1, 2$. If the condition does hold for *all* $t \in [0, 1]$, then we fail to reject the null hypothesis at level α. In contrast, if the parametric fit of the function is not *entirely covered* by the constructed UCB, then the null is rejected.

The advantage of the UCB over the traditional point-wise bands is that the UCB can be used for simultaneous inference of an unknown function by allowing us to figure out the overall shape of the function. In addition, the UCB is a more conservative confidence band than the point-wise ones in that the UCB is wider than its point-wise counterpart. Thus, any test results based on the UCB would be more robust than those under the point-wise ones. For these reasons, the UCB has recently gained more attention in econometrics and statistics literature. For example, Zhou and Wu (2010) show how to construct the UCBs of time-varying regression coefficients and apply the method for the Hong Kong circulatory and respiratory data. Kim (2013) applies the idea of UCB to the semi-parametric environmental Kuznets curve (EKC) and test the hypothesis of an inverted U-shaped pattern for the income-pollution relationship. The paper finds out that the standard parametric forms for the EKC are rejected by the UCB. Furthermore, Kim (2014b) constructs the UCB of the long-run trend in unemployment rate, typically known as the $NAIRU$ parameter of the Phillips Curve, and conduct simultaneous inference of the structural parameter to test for its uniform constancy over the years.

Given these various interesting results, we shall construct the UCB of the slope coefficient (6), and carry out inference. To our best knowledge, this is the first time that one performs a UCB-based test on parametric specifications of the slope coefficient in the forward premium regression. Technical details of constructing the UCB are summarized in Appendix 2 of this paper.

4 Empirical results

The data used in this work are spot and one month forward exchange rate data involving the nine currencies of Australian Dollar (AUD), Canadian Dollar (CAD), Swiss Franc (CHF), Danish Krone (DKK), British Pound (GBP), Japanese Yen (JPY), Norwegian Krone (NOK), New Zealand Dollar (NZD), vis a vis U S Dollar (USD), [4] which is the numeraire currency in our study. This study uses end-of-month observations from December 1988 through October 2010. A major reason for using monthly data is to match the ex ante returns from UIP with fundamentals that are only observable at the monthly frequency. This is also consistent with most of the models for risk premium that emphasize

[4]The Euro that started in January 1999 is not used in this study to ensure that our sample is long enough for the non-parametric estimation and inference.

the volatility of fundamentals[5] . Also, the use of monthly data makes rational expectations errors from *UIP uncorrelated*, which is a statistical convenience.

The estimates of the time-varying slope coefficient and their corresponding 95% UCBs for the various currencies, are reported in Figures 1 and 2. The solid curves represent the local-linear regression estimates of the slope coefficients in (7), while the dashed bands are the 95% UCBs of the parameters. The optimal bandwidths for the local-linear regression chosen by the GCV are also reported in the figures. The solid horizontal lines represent the hypothesis that $H_0 : \beta_1(\cdot) = 1$. Figures 1 and 2 also present the OLS estimates of fixed β_1 in (5) for reference. In addition, Table 3 reports the proportion of $\hat{\beta}_1(t) > 1$ for each currency, where $\hat{\beta}_1(t)$ is the local-linear regression estimate of $\beta_1(t)$ in (6).

As seen from Figures 1 and 2, there is a substantial amount of time variation in the slope coefficients for all eight currencies, which are generally found to take both positive and negative values across time. Except for the CAD and GBP, the slope coefficients stay predominantly negative throughout the 1990s, which is consistent with studies such as Baillie and Bollerslev (2000) and Lothian and Wu (2011). However, there is some variation across currencies with the CAD having a positive slope coefficient during the second half of the 1990s; and also for the GBP which is briefly positive around 1995.

There is also considerable evidence that the slope coefficients for all eight currencies tend to move together (see Table 2) and that the coefficients for all eight currencies, except for DKK, take *positive* values during from about 2007 onwards, which corresponds to the severe financial crisis. The slope coefficient for the AUD, CAD, CHF, JPY, NOK and NZD, becomes increasingly positive during the late 2000s, while the slope coefficient for the GBP is only barely positive during the period. This distinctive co-movements among these slope coefficients (see Table 2) is very pronounced during the crisis and highly

[5]A counter example is a model for risk premium based on central bank intervention as in Baillie and Osterberg (1997), and which correspondingly used daily data.

suggestive of common factors determining their movements. Furthermore, after 2010, the slope coefficients for the DKK and GBP turn downward, while the coefficients for the other six currencies continue to climb upward.

From Fama (1984), it is well-known that the slope coefficient of forward premium regression (5) can be expressed

$$\beta_1 = \frac{Cov\left(p_\tau,\, q_\tau\right) + Var\left(q_\tau\right)}{Var\left(p_\tau + q_\tau\right)}$$

where $p_\tau = f_\tau - \mathbb{E}\left(s_{\tau+1}|\mathcal{I}_\tau\right)$ is the *risk premium* and $q_\tau = \mathbb{E}\left(s_{\tau+1}|\mathcal{I}_\tau\right) - s_\tau$ is the *expected rate of depreciation*. Here $\mathcal{I}_\tau$ is the information set available up to time τ. Fama (1984) shows that $\beta_1 < 0$ requires both $Cov\left(p_\tau,\, q_\tau\right) < 0$ and $Var\left(p_\tau\right) > Var\left(q_\tau\right)$. That is, if $\beta_1(t) < 0$ for some t, then both of these conditions are satisfied during that period. In contrast, if $\beta_1(t) > 0$, then at least one of these conditions are not satisfied. Figures 1 and 2 shows that throughout the financial crisis of late 2000s, we have either $Cov\left(p_\tau,\, q_\tau\right) > 0$ or $Var\left(p_\tau\right) < Var\left(q_\tau\right)$ for all the currencies, except for the DKK. Typically, investors demand higher premium to keep holding depreciating currencies during financially unstable periods such as the late 2000s, which leads to the positive co-movement between the risk premium and expected depreciation during the period. This basic intuition appears to be well illustrated in Figures 1 and 2 by the increasingly positive β_1's from the late 2000s.

4.1 Testing the UIP hypothesis

As previously explained, a great attraction of using the UCB of $\beta_1(\cdot)$ in this study is to carry out simultaneous inference of the unknown parameter function, and testing any hypothesis is simply facilitated by checking whether the specification is *contained* by the UCB. From Figures 1 and 2, it is easily seen that the null hypothesis is *rejected* at 5% for all eight currencies because the solid horizontal line (*i.e.* $\beta_1 = 1$) cannot be entirely contained by the 95% UCBs. In fact, even more general hypothesis $H_0 :\ \beta_1(\cdot) = c$, where

c is some constant, is also rejected at 5% for all the currencies since no horizontal line can be placed entirely within the constructed 95% UCBs.

The Local Deviation from Uncovered Interest Parity ($LDUIP$) is given by

$$y_{\tau+1} = \Delta s_{\tau+1} - \widehat{\beta}_0 \left(\frac{\tau}{T}\right) - \widehat{\beta}_1 \left(\frac{\tau}{T}\right) (f_\tau - s_\tau)$$

and the percentage of the number of times that $\beta_1(\cdot) = 1$ is not covered or included by the UCB, is presented in Table 4. Moreover, Table 3 illustrates the proportion of $\hat{\beta}_1(\cdot)$ greater than one. These results indicate that the percentage of occasions when $\beta_1(\cdot) > 1$ is 75% for Canada and around 40% for the UK, Norway and New Zealand. Table 4 shows corresponding percentages for when the direct UIP hypothesis $\beta_1(\cdot) = 1$ is violated. The hypothesis is rejected for 62% of the time periods for the Australian \$, and over 30% of the time for all currencies except Canada and Switzerland. Hence the empirical results confirm the forward premium anomaly for all the currencies for many of the time periods.

The key reason for the rejection of the null $H_0 : \beta_1(\cdot) = c$ is that the slope coefficient changes dramatically after the mid 2000s. For the AUD, CAD, CHF, JPY, NOK and NZD, the coefficients increase significantly during this period. On the other hand, β_1 for the DKK and GBP first increases during the early/mid 2000s, and then decreases, dramatically for the DKK and mildly for the GBP, during the late 2000s, which leads to the rejection of the null. Hence the approach in this study indicates the rejection of UIP for many time periods, and an additional attraction of the procedure is that there is clear co-movement among the slope coefficients for the different currencies. Large negative coefficients in the 1980s and 1990s are replaced, for many currencies, with a reversal of the forward premium anomaly after the financial crisis of 2008. This emphasizes the substantial variation in the slope coefficient and the inappropriateness of asserting that the anomaly exists for all time periods.

4.2 Determinants of the time varying beta: role of risk and fundamentals

Articles by Hansen and Hodrick (1983), Domowitz and Hakkio (1985), Hodrick (1987, 1989), Bekaert and Hodrick (1993), Baillie and Osterberg (1997), Mark and Wu (2000) and Verdelhan (2012) have all provided detailed models of time-dependent risk premium, which have had variable degrees of empirical success. However, there has been no clear and definitive model of time-varying risk that has been found to be reliable across currencies and different time periods. The relatively smooth estimates of $\beta_1(\cdot)$ obtained by the kernel weighted regression show slow changes that turn out to be highly predictable from quite conventional models of time-varying risk premium and also asset model fundamentals. Given $\hat{\beta}_1(\cdot)$ from (6), the θ in (10) can be estimated by *OLS* and then a *Wald test* is used to test the hypothesis that $H_0 : \theta = 0$ versus $H_1 : H_0$ is incorrect. Hence a rejection of the hypothesis indicates that the violations of *UIP* can be at least partly explained by standard fundamentals and measures of risk. The following regression is then estimated,

$$\hat{\beta}_1 \left(\frac{\tau}{T} \right) - 1 = \mathbf{x}_{\tau-1}^{/} \theta + \xi_\tau \tag{10}$$

where θ is a parameter vector and $\hat{\beta}_1(\cdot)$ is the non-parametric estimate of $\beta_1(\cdot)$ in (6). Here ξ_τ is a mean zero error that is uncorrelated with $\mathbf{x}_{\tau-1}$. The covariates $\mathbf{x}_{\tau-1}$ in (10) are the following menu:

$$\mathbf{x}_{\tau-1}^{/} = \left[\Delta m_{\tau-1}, \Delta m_{\tau-1}^{*}, \Delta y_{\tau-1}, \Delta y_{\tau-1}^{*}, \left(i_{\tau-1}^{2} - i_{\tau-1}^{*2} \right), Var_{\tau-2} \Delta m_{\tau-1} \right] \tag{11}$$

where $\Delta m_{\tau-1}$ is the change in the log of the *US* money supply, $\Delta y_{\tau-1}$ is the change in the log of the *US* index of industrial production, $(*)$ denotes *foreign* equivalents, while $\left(i_{\tau-1}^{2} - i_{\tau-1}^{*2} \right)$ is the differential between squared nominal 30-day *T Bill* interest rates for the relative volatility in the two bond markets; and $Var_{\tau-2} \Delta m_{\tau-1}$ represents the conditional variance of *US* money growth rates. This last variable was generated from a *GARCH* model, and is used following the findings in Hodrick (1989) and in Baillie and Kilic (2006).

The Wald test statistics presented in Table 7 reveal substantial predictability of $\hat{\beta}_1(\cdot)$ from information in the lagged fundamentals and risk premium variables in (11). Similar analysis only based on information twelve months previously, as opposed to one month ago, is presented in Table 8. Again, the Wald test overwhelmingly rejects the null of no significance of the fundamentals for all cases for all eight currencies. The R-squared from the regressions reported in Tables 5–8 ranges from 15% to 35% over the eight currencies being considered and indicate a substantial amount of predictability in the movement of the local $\hat{\beta}_1(\cdot)$ coefficient[6].

The regression parameter estimates corresponding to the Wald tests in Tables 7 and 8 for equation (10) are reported in Tables 5 and 6, respectively. The impact of US money growth rates, and also the volatility of US money growth rates are seen to be generally statistically significant and positive for most of the currencies. The exception is for the Danish Krone, which moves in the opposite and non-intuitive direction. Overall, the effects also indicate substantial non-linearities with shocks on US money growth rates and its associated volatility leading to proportional impacts of the interest rate differential on spot exchange rate returns.

Given that the $LDUIP$ are in some sense "model-free", deterministic estimates of the slope parameter in the UIP regression, interest focuses on the reasons for the time variation in these coefficients. Various fundamental-based explanations and also models based on some models developed to explain time-dependent risk premia are used in a second step analysis, where the generated deterministic β_t are regressed on four different

[6]The inclusion of these variables in the standard rejection equation (10) would generally not lead to significant results due to the relatively very high volatility of spot returns tendings to dwarf the far smaller movements in the fundamentals. Hence our two-step analysis of the estimated beta slope coefficient has more economic interpretation.

risk premium models, which typically contain estimated second moments, or conditional variances and covariances of some variables associated with previously developed economic models of risk premium. The validity and relative strength of each model is then assessed through a classical frequentist-based model averaging procedure. This method indicates the most likely reason for the breakdown of UIP over the whole sample and also for certain sub-periods such as the financial crisis. To this end, the paper employs the *model averaging* technique. Hjort and Claeskens (2003) provide an excellent review of various model averaging techniques.

Specifically, we implement the least squares model averaging ($LSMA$) approach (Hansen, 2007). While existing model average methods are typically based on exponential Akaike information criterion (AIC) (Akaike, 1979) and Bayesian information criterion (BIC) (Schwarz, 1978), the $LSMA$ proposes selecting the weights by minimizing a $Mallows$ criterion (Mallows, 1973), an estimate of the average squared error from the model average fit. This new Mallows model average (MMA) estimator is asymptotically optimal in the sense that it achieves the lowest possible squared error in a class of discrete model average estimators (Hansen, 2007). In addition, this approach is free of any distributional assumption on the model error. In various simulation experiments, the MMA estimator compares favorably to those based on AIC and BIC weights (Hansen, 2007).

Several different models are used to represent the time-dependent risk premium, or mis-specification of the UIP relationship. The first model is that

$$\hat{\beta}_1\left(\frac{\tau}{T}\right) - 1 = \theta_{1,1}\Delta m_{\tau-1} + \theta_{1,2}\Delta m^*_{\tau-1} + \theta_{1,3}\Delta y_{\tau-1} + \theta_{1,4}\Delta y^*_{\tau-1} + \xi_{1,\tau} \tag{12}$$

where Δm_τ is the change in the log of the US money supply, Δy_τ is the change in the log of the US index of industrial production, $(*)$ denotes *foreign* equivalents. This model to explain movements in $\hat{\beta}_1(\cdot)$ simply means that changes in the ex ante returns over UIP partially explicable in terms of changes in fundamentals from the monetary model. There are several potential economic interpretations of this which would appear to imply some form of mis-pricing of the UIP is 1 have been omitted from the UIP relationship.

However, the principal interest in this study is the notion of risk and to this end a seminal paper on this topic was by Domowitz and Hakkio (1985) who developed a theoretical two-country-, two-good-type model with specific endowments and Cobb Douglas utility functions of the agents to derive a risk premium of

$$\hat{\beta}_1\left(\frac{\tau}{T}\right) - 1 = \theta_{2,1} Var_{\tau-2}\left(\Delta m_{\tau-1}\right) + \theta_{2,2} Var_{\tau-2}\left(\Delta m^*_{\tau-1}\right) + \xi_{2,\tau} \tag{13}$$

where there is an additional restriction that $\theta_{2,1} = -\theta_{2,2} = 0.5$. Baillie and Kilic (2006) find evidence that the volatility of US money growth rates is important in a logistic smooth transition auto-regression for representing transitions between regimes where UIP holds or is violated. A further model due to Giovannini and Jorion (1987) that postulates the differential volatility in the bond market, is

$$\hat{\beta}_1\left(\frac{\tau}{T}\right) - 1 = \theta_{3,1}\left(i^2_{\tau-1} - i^{*2}_{\tau-1}\right) + \xi_{3,\tau} \tag{14}$$

A more complete model for time-dependent risk premia in the foreign exchange market is due to Hodrick (1989) who uses an overlapping-generations model with cash-in-advance constraints to derive the risk premium as,

$$\hat{\beta}_1\left(\frac{\tau}{T}\right) - 1 = \theta_{4,1} Var_{\tau-2}\left(\Delta m_{\tau-1}\right) + \theta_{4,2} Var_{\tau-2}\left(\Delta m^*_{\tau-1}\right) + \theta_{4,3} Var_{\tau-2}\left(\Delta y_{\tau-1}\right) + \theta_{4,4} Var_{\tau-2}\left(\Delta y^*_{\tau-1}\right) + \xi_{4,\tau} \tag{15}$$

Since the volatility of the fundamentals is unobservable, they need to be generated from uni-variate $ARMA-GARCH$ models.

The root-mean-squared-errors ($RMSE$) under (12)–(15) and the $RMSE$ under the $LSMA$ of (12)–(15) are reported by Table 9. For the sake of comparison, each $RMSE$ is divided by the lowest $RMSE$ for the corresponding currency. Hence the best performance for each currency is represented by unit and is shown *in bold*. As Table 9 clearly shows it, the $LSMA$ outperforms the individual risk premium models for most currencies. The only exception is the case of Danish Krone (DKK). However, the difference in $RMSE$ between the best one and the model-averaging one is almost negligible.

5 Conclusion

This paper has used recently developed kernel smoothing regression procedures to derive uniform confidence bounds to investigate the forward premium anomaly, where spot currency returns are generally found to be negatively correlated with lagged interest rate differentials, or forward premium. The econometric techniques used in this paper have considerable advantages over simple rolling regression methods and they also provide relatively tight confidence intervals. The results indicate remarkable variation in the time periods where the anomaly occurs and where the deviatons from uncovered interest rate parity (UIP) are relatively small and fall within the bands of UIP. There is also some considerable similarity in *co-movements across currencies*. Hence, contrary to the established beliefs, the anomaly does not hold continuously and there are many time periods where the hypothesis of UIP cannot be rejected. The departures from UIP throughout the sample were found to be partly predictable and to be based on many of the standard fundamentals associated with the monetary model, and also variables associated with time-dependent risk premium.

As previously noted, the traditional models for time-dependent risk premium and other explanations of the forward premium anomaly have not been very empirically successful. The results obtained in this study, suggest that there is exploitable information in a function of the fundamentals and possible risk terms, which allow predictability of the extent, and even degree of persistence, of the anomaly. An interesting issue for future research is whether any existing models, or *combination of models,* might be consistent with this function of information, and how the success of the models varies over time. A full investigation of these issues is intriguing and is the subject of future research of the authors.

6 Appendix 1: Assumptions of Kernel Smoothing

First, we introduce notations. For any vector $\mathbf{v} = (v_1, v_2, \ldots, v_p) \in \mathbb{R}^p$, we let $|\mathbf{v}| = \left(\sum_{i=1}^p v_i^2\right)^{1/2}$. For any random vector $\mathbf{V}$, we write $\mathbf{V} \in \mathcal{L}^q$ $(q > 0)$ if $\|\mathbf{V}\|_q = \left[\mathbb{E}\left(|\mathbf{V}|^q\right)\right]^{1/q} < \infty$. In particular, $\|\mathbf{V}\| = \|\mathbf{V}\|_2$. We denote $\mathbf{L} : [0,1] \times \mathbb{R}^\infty \mapsto \mathbb{R}^p$ as a measurable function such that $\mathbf{L}(t, \mathcal{F}_i)$ is a properly defined $(p \times 1)$ random vector for all $t \in [0,1]$, where $\mathcal{F}_i = (\cdots, \eta_{i-1}, \eta_i)$ with independent and identically distributed (IID) random errors $\{\eta_j\}_{j\in\mathbb{Z}}$. Define the physical dependence measure (Wu, 2005) for $\mathbf{L}(t, \mathcal{F}_i)$ as the following:

$$\delta_q(\mathbf{L}, k) = \sup_{t\in[0,1]} \|\mathbf{L}(t, \mathcal{F}_k) - \mathbf{L}(t, \mathcal{F}_k^*)\|_q \tag{16}$$

where $\mathcal{F}_i^* = (\cdots, \eta_0^*, \cdots, \eta_{i-1}, \eta_i)$ is a coupled process with η_0^* being an IID copy of η_0. For discussion on this dependence measure, we refer to Wu (2005). For a class of stochastic processes $\{\mathbf{L}(t, \mathcal{F}_i)\}_{i\in\mathbb{Z}}$, we say that the process is $\mathcal{L}^q$ *stochastic Lipschitz-continuous* over $[0,1]$ if the following condition holds:

$$\sup_{0\le t_1 < t_2 \le 1} \frac{\|\mathbf{L}(t_2, \mathcal{F}_0) - \mathbf{L}(t_1, \mathcal{F}_0)\|_q}{|t_2 - t_1|} < \infty \tag{17}$$

We denote a collection of such systems by Lip_q. The required assumptions are:

Assumption 1. *Let covariates $\mathbf{x}_\tau$ of (7) be*

$$\mathbf{x}_\tau = \mathbf{G}(\tau/T, \mathcal{U}_\tau)$$

where $\mathbf{G} := (G_1, \cdots, G_p)^\top$ is a measurable function such that $\mathbf{G}(t,\mathcal{U}_\tau)$ is well-defined for each $t \in [0,1]$. Here $\mathcal{U}_\tau = (\cdots, u_{\tau-1}, u_\tau)$ with IID errors $\{u_j\}_{j\in\mathbb{Z}}$. Moreover, $\mathbf{G}(t,\ \mathcal{U}_\tau) \in$

Lip$_2$ *and* $\sup_{0\leq t\leq 1} \|\mathbf{G}(t,\ \mathcal{U}_\tau)\|_4 < \infty$.

It should be noted that the Assumption 1 allows the regressors $\mathbf{x}_\tau$ to be *non-stationary* since their moments are allowed to be time-varying. Generally, the time variation in these characteristics is assumed to be smooth, rather than abrupt. Note also that $\mathbf{x}_\tau$ is dependent due to the cumulative IID random elements. Specifically, Assumption 1 ensures that model regressors $\mathbf{x}_\tau$ are *locally stationary*, which is a mild form of non-stationarity. That is, if one observes locally stationary variables in a relatively short time span, they are approximately stationary. However, in the long run, the variables behave as non-stationary ones. Since many economic variables are possibly locally stationary processes, we can make our model specification more general by introducing this assumption. For more on the local-stationarity, we refer to Ombao, Von Sachs & Guo (2005), Zhou & Wu (2010) and Kim (2014a, 2014b).

Assumption 2. *Let* $M(t) := \mathbb{E}\left[\mathbf{G}(t,\mathcal{U}_0)\mathbf{G}(t,\mathcal{U}_0)^\top\right]$. *Then, the smallest eigenvalue of* $M(t)$ *is bounded away from* 0 *on* $[0,1]$.

This assumption prevents the asymptotic multicollinearity of regressors.

Assumption 3. *The error* ϵ_τ *of (6) forms a stationary martingale-difference process such that*

$$\epsilon_\tau = H(\mathcal{V}_\tau)$$

where $H(\cdot)$ *is a measurable function and* $\mathcal{V}_\tau = (\cdots, v_{\tau-1}, v_\tau)$ *with IID errors* $\{v_j\}_{j\in\mathbb{Z}}$. *Here* $\mathbb{E}(\epsilon_\tau|\mathbf{x}_\tau) = 0$.

Note that this assumption allows the error term to be *dependent* but *uncorrelated*. The dependence structure for ϵ_τ is flexible and general in that function $H(\cdot)$ is not specified. For example, a stationary $ARCH$ process (Engle, 1982) would satisfy these requirements. Also, this assumption ensures that the error is uncorrelated with the regressors.

Assumption 4. *Let* $\mathbf{U}(t, \mathcal{I}_\tau) = \mathbf{G}(t, \mathcal{U}_\tau)H(\mathcal{V}_\tau)$ *where* $\mathcal{I}_\tau = (\cdots, \zeta_{\tau-1}, \zeta_\tau)$ *and* $\zeta_j = (u_j, v_j)^{/}$. *Define*

$$\Lambda(t) := \sum_{k \in \mathbb{Z}} cov\left(\mathbf{U}(t, \mathcal{I}_0),\, \mathbf{U}(t, \mathcal{I}_k)\right)$$

where the smallest eigenvalue of $\Lambda(t)$ *is bounded away from* 0 *on* $[0, 1]$.

Assumption 5. *Let* $\sum_{\ell=0}^{\infty}[\delta_4(\mathbf{G}, \ell) + \delta_2(\mathbf{U}, \ell)] < \infty$.

This assumption also ensures short-range dependence among the variables in our model. The interpretation is that the cumulative effect of a single error on all future values is bounded. The measure of dependence used here is the physical dependence measure (Wu, 2005) based on causal processes. This measure is known to be particularly useful for characterizing dependence in non-linear time series models (Wu, 2005; Zhou & Wu, 2010; Kim, 2014a).

Assumption 6. *The non-parametric functions* $\beta_0(\cdot)$ *and* $\beta_1(\cdot)$ *are twice continuously differentiable over the compact domain* $[0, 1]$.

This guarantees that the parameter functions $\beta_0(\cdot)$ and $\beta_1(\cdot)$ change smoothly over time. In particular, the second-order continuity of the parameters is required for the weak consistency of the local-linear estimates of their first-order derivatives.

Assumption 7. *Let the kernel function* $K(\cdot)$ *be bounded, symmetric, with bounded support* $[-A, A]$, $K \in \mathcal{C}^1[-A, A]$, $K(\pm A) = 0$ *and* $\sup_u |K'(u)| < \infty$.

This allows popular kernel functions such as the Epanechnikov kernel, which is used in the non-parametric estimation of $\beta_0(\cdot)$ and $\beta_1(\cdot)$.

7 Appendix 2: Construction of the UCB

Recall the bias-corrected estimator $\tilde{\boldsymbol{\beta}}(t)$, given by (8). Under Assumptions 1–7 and $Tb^7\log(T) + \frac{(\log(T))^3}{bT^{2/5}} = o(1)$, Theorem 3 in Zhou and Wu (2010) shows:

$$\mathbb{P}\left\{\sqrt{\frac{Tb}{\lambda_K}}\sup_{t\in[0,1]}\left|\Sigma^{-1}(t)\left(\tilde{\boldsymbol{\beta}}(t)-\boldsymbol{\beta}(t)\right)\right| - d_T \leq \frac{u}{\sqrt{2\log(b^{-1})}}\right\} \quad \to \quad e^{-2e^{-u}} \tag{18}$$

where $\lambda_K = \int_{\mathbb{R}} K^2(u)du$ and $\Sigma^2(t) = M^{-1}(t)\cdot\Lambda(t)\cdot M^{-1}(t)$ under Assumptions 2 and 4. Here the centering parameter d_T is given by:

$$d_T = \sqrt{2\log(b^{-1})} + \frac{1/2\log\left(\log\left(b^{-1}\right)\right) + 1/2\log\left(\int_{\mathbb{R}}(K'(u))^2du/\left(4\pi\lambda_K\right)\right)}{\sqrt{2\log(b^{-1})}}$$

Note that the convergence rate to the asymptotic Gumbel distribution in (18) is $1/\sqrt{\log(T)}$, since $b \asymp T^{-\alpha},\ \alpha\in(0,1)$. Given this slow rate of convergence, we employ the following invariance principle (Zhou and Wu, 2010):

$$\sqrt{\frac{Tb}{\log(T)}}\sup_{t\in[0,1]}\left|\tilde{\boldsymbol{\beta}}(t)-\boldsymbol{\beta}(t)-\Sigma(t)\sum_{\tau=1}^{T} w_T^*(t,\tau)\mathbf{Z}_\tau\right| = o_{\mathbb{P}}(1) \tag{19}$$

where $\mathbf{Z}_\tau \sim NID\left(\mathbf{0},\mathbf{I}d_2\right)$ and $w_T^*(t,\tau) = \frac{1}{Tb}K^*\left(\frac{t-\tau/T}{b}\right)$ with $K^*(u) = 2\sqrt{2}K(\sqrt{2}u) - K(u)$ from the bias-correction. Here $\Sigma(t)$ is introduced by (18). Then, by (19), it is then possible to construct the UCB of the slope coefficient $\beta_1(t)$, $0\leq t\leq 1$, in (6). The procedure requires the following steps:

(i) (Bandwidth selection) Consider $\tilde{\boldsymbol{\beta}}(t)$ in (8) under some b and the Epanechnikov kernel. Then, the fitted values for (6) would be $\widehat{\Delta s_{\tau+1}}(b) = \tilde{\beta}_0(\tau/T)+\tilde{\beta}_1(\tau/T)(f_\tau - s_\tau),\ \ \tau =$

$1, \cdots, T-1$. Note here that $\widehat{\Delta s_{\tau+1}}(b)$ depends on b due to $\tilde{\beta}_0(\tau/T)$ and $\tilde{\beta}_1(\tau/T)$. To pick up the *optimal* bandwidth, we consider:

$$S(b) = H(b)Y,$$

where $S(b) := \left(\widehat{\Delta s_1}(b), \widehat{\Delta s_2}(b), \cdots, \widehat{\Delta s_T}(b)\right)'$, $Y := (\Delta s_1, \cdots, \Delta s_T)'$ and $H(b)$ is a $(T \times T)$ smoothing matrix that depends on b. The GCV criterion (Craven and Wahba, 1979) chooses the optimal bandwidth b^{opt} that minimizes the following criterion:

$$b^{opt} := \underset{b}{\operatorname{argmin}} \frac{T^{-1} \sum_{\tau=1}^{T} \left(\Delta s_\tau - \widehat{\Delta s_\tau}(b)\right)^2}{\{1 - \operatorname{trace}[H(b)]/T\}^2}$$

where the numerator represents the goodness-of-fit and the denominator can be viewed as the model's degrees of freedom (Kim, 2014b). We obtain (8) using b^{opt}.

(ii) Compute $\sup_{0 \le t \le 1} \left| \sum_{\tau=1}^{T} w_T^*(t,\tau) Z_\tau \right|$, where $\{Z_\tau\}$ are generated as $NID(0,1)$ random variables and $w_T^*(t,\tau) = \frac{1}{Tb} K^* \left(\frac{t-\tau/T}{b}\right)$ with higher-order kernel $K^*(u) = 2\sqrt{2}K(\sqrt{2}u) - K(u)$.

(iii) Repeat (ii), say $1,000$ times, and then obtain the 95th quantile of the sampling distribution of $\sup_{0 \le t \le 1} \left| \sum_{\tau=1}^{T} w_T^*(t,\tau) Z_\tau \right|$, and denote it as $\hat{q}_{0.95}$.

(iv) Estimate $\Sigma^2(t)$ by:

$$\hat{\Sigma}^2(t) := \hat{M}^{-1}(t) \cdot \hat{\Lambda}(t) \cdot \hat{M}^{-1}(t)$$

where $\hat{M}(t)$ and $\hat{\Lambda}(t)$ are the estimates of $M(t)$ and $\Lambda(t)$ in Assumptions 2 and 4. The estimates are provided by Section 4.3 of Zhou and Wu (2010). Denote the (i,j)th element of $\widehat{\Sigma}^2(t)$ by $\hat{\sigma}_{i,j}^2(t)$.

(v) The 95% UCB of $\beta_1(t)$ is $\left[\tilde{\beta}_1(t) \pm \hat{q}_{0.95} \hat{\sigma}_{2,2}(t)\right]$.

References

[1] Akaike, H. (1979), A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika* 66, 237-242.

[2] Baillie, R.T. and Bollerslev, T. (1994). The long memory of the forward premium. *Journal of International Money and Finance* **13**, 565–571.

[3] Baillie, R.T. and Osterberg, W.P. (1997). Central bank intervention and risk in the forward market. *Journal of International Economics* **43**, 483–497.

[4] Baillie, R.T. and Bollerslev, T. (2000). The forward premium anomaly is not as bad as you think. *Journal of International Money and Finance* **19**, 471–488.

[5] Baillie, R.T. and Kilic, R. (2006). Do asymmetric and nonlinear adjustments explain the forward premium anomaly? *Journal of International Money and Finance* **25**, 22–47.

[6] Baillie, R.T., Lippens, R.E. and McMahon, P.C. (1983). Testing rational expectations and efficiency in the foreign exchange market. *Econometrica* **51**, 553–563.

[7] Bansal, R. (1997). An exploration of the forward premium puzzle in currency markets. *Review of Financial Studies* **10**, 369–403.

[8] Bansal, R. and Dahlquist, M. (2000). The forward premium puzzle: different tales from developed and emerging economies. *Journal of International Economics* **51**, 115–144.

[9] Bekaert, G. and Hodrick, R.J. (1993). On Biases in the measurement of foreign exchange risk premiums. *Journal of International Money and Finance* **12**, 115–138.

[10] Chang, S.S. (2013). Can cross-country rebalancing give rise to forward bias in FX markets. *Journal of International Money and Finance* **32**, 1079-1096.

[11] Chen, R. and Tsay, R.S. (1993). Functional-coefficient autoegressive models. *Journal of the American Statistical Association* **88**, 298–308.

[12] Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterlpots. *Journal of the American Statistical Association* **74**, 829–836.

[13] Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions. *Numerische Mathematik* **31**, 377–403.

[14] Domowitz, I. and Hakkio, C.S. (1985). Conditional variance and the risk in the foreign exchange market. *Journal of International Economics* **19**, 47–66.

[15] Engel, C. and K.D. West (2005). Exchange rates and fundamentals. *Journal of Political Economy* **113**, 485-517.

[16] Engle, R.F (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987-1007.

[17] Fama E. (1984). Forward and spot exchange rates. *Journal of Monetary Economics* **14**, 319–338.

[18] Froot, R.A. and Thaler, R.H. (1990). Anomalies: foreign exchange. *Journal of Economic Perspectives* **4**, 179–192.

[19] Giovannini, A. and Jorion, P. (1987). Interest rates and risk premium in the foreign exchange and stock market. *Journal of International Money and Finance* 6, 107-123.

[20] Hakkio, C.S. (1981). Expectations and the forward exchange rate. *International Economic Review* **22**, 663–678.

[21] Hansen, B.E. (2007). Least squares model averaging. *Econometrica* **75**, 1175–1189.

[22] Hansen, L.P. and Hodrick, R.J. (1980). Forward exchange rates as optimal predictors of future spot rates: an econometric analysis. *Journal of Political Economy* **88**, 829–853. Jacob A. Frenkel.

[23] Hansen, L.P. and Hodrick, R.J. (1983). Risk averse speculation in the forward foreign exchange market: an econometric analysis of linear models. *Exchange Rates and International Economics*, edited by Jacob A. Frenkel.

[24] Hodrick, R.J. (1987). *The empirical evidence on the efficiency of forward and futures foreign exchange markets*. Harwood, London.

[25] Hodrick, R.J. (1989). Risk, uncertainty and exchange rates. *Journal of Monetary Economics* **23**, 433–459.

[26] Hjort, N.L. and Claeskens, G. (2003), Frequentist model average estimators, *Journal of the American Statistical Association,* 98, 879-899.

[27] Kim, K.H. (2013) Inference of the environmental Kuznets curve. *Applied Economics Letters* **20**, 119–122.

[28] Kim, K.H. (2014a) Counter-cyclical risk aversion. *Journal of Empirical Finance* **29**, 384–401.

[29] Kim, K.H. (2014b) Inference of the trend in a partially linear model with locally stationary regressors. *Econometric Reviews*, DOI: 10.1080/07474938.2014.976530.

[30] Lothian, J.R. and Wu, L. (2011). Uncovered interest rate parity over the past two centuries. *Journal of International Money and Finance* **30**, 448–473.

[31] Mallows, C.L. (1973). Some comments on C_p. *Technometrics* **15**, 661–675.

[32] Mark, N.C. and Wu, Y. (1997). Rethinking deviations from uncovered interest rate parity: the role of covariance risk and noise. *Economic Journal* **108**, 1686–1706.

[33] Maynard, A. and Phillips, P.C.B. (2001). Rethinking old empirical puzzle: econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics* **16**, 671–708.

[34] Ombao, H., Von Sachs, R. and Guo, W. (2005). SLEX analysis of multivariate non-stationary time series. *Journal of American Statistical Association* **100** 519–531.

[35] Orbe, S., Ferreira, E. and Rodrigues-Poo, J., (2005). Nonparametric estimation of time varying parameters under shape restrictions. *Journal of Econometrics* **126**, 53–77.

[36] Phillips, P.C.B. (2001). Trending time series and macroeconomic activity: some present and future challenges. *Journal of Econometrics* **100**, 21–27.

[37] Sakoulis, G., Zivot, E. and Choi, K. (2010). Time variation and structural change in the forward discount: implications for the forward rate unbiasedness hypothesis. *Journal of Empirical Finance*, **17**, 957–966.

[38] Schwarz, G.E. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464.

[39] Wolf, C.C.P. (1987). Time-varying parameters and the out-of-sample forecasting performance of structural exchange rate models. *Journal of Business and Economic Statistics* **5**, 87–97.

[40] Wu, W.B. (2005). Nonlinear system theory: another look at dependence. *Proceedings of the National Academy of Sciences USA*. **102**, 14150–14154.

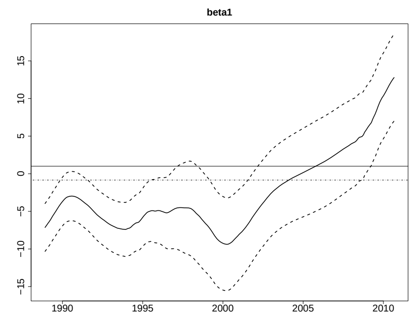
[41] Zhou, Z. and Wu, W.B. (2010). Simultaneous inference of linear models with time varying coefficients. *Journal of the Royal Statistical Society, Series B*. **72**, 513–531.

Table 1: Mean of time-varying beta

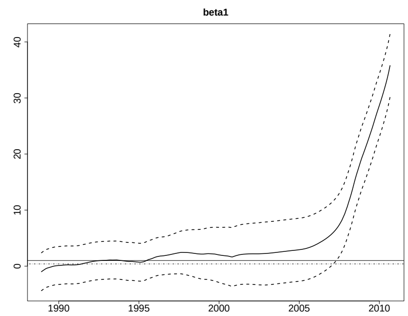
currency	mean of $\beta_1(t)$
AUD	-2.5899
CAD	4.6792
CHF	0.4787
DKK	-2.5047
GBP	-0.5552
JPY	-1.9178
NOK	2.4767
NZD	3.7606

Table 2: Correlation among time-varying betas

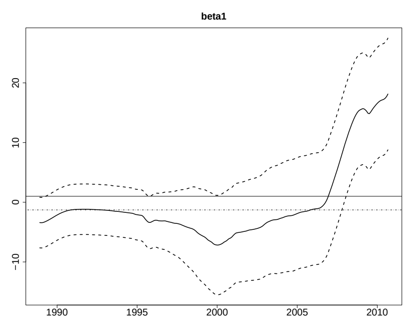
	AUD	CAD	CHF	DKK	GBP	JPY	NOK	NZD
AUD	1	0.8621	0.8745	-0.0204	0.1797	0.8390	0.8530	0.9281
CAD		1	0.9667	-0.1527	0.2287	0.9054	0.9661	0.9308
CHF			1	0.0496	0.2613	0.8541	0.9835	0.9262
DKK				1	0.6280	-0.2816	0.0889	0.0792
GBP					1	-0.0073	0.3798	0.4364
JPY						1	0.8239	0.8507
NOK							1	0.9496
NZD								1



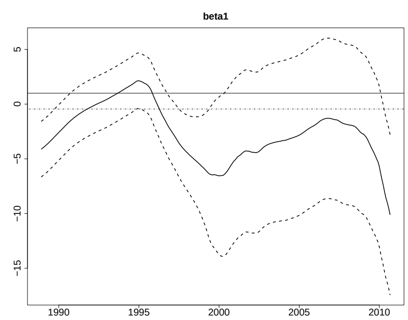
(a) Australian Dollar (AUD)



(b) Canadian Dollar (CAD)

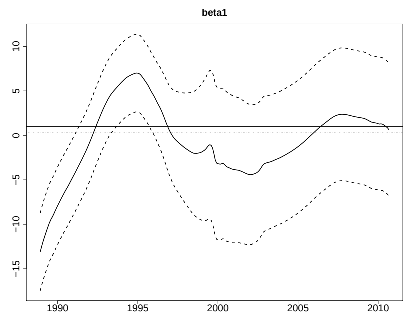


(c) Swiss Franc (CHF)

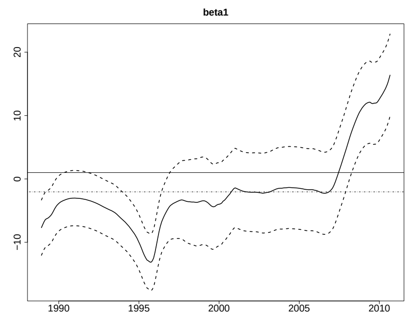


(d) Danish Krone (DKK)

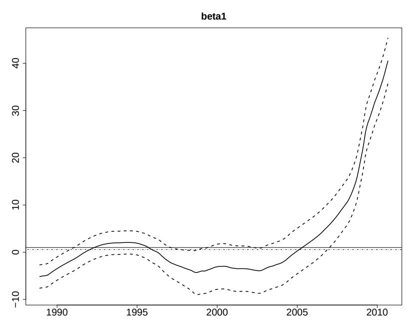
Figure 1: The *solid* curve is the local-linear regression estimate of $\beta_1(\cdot)$ in (6). The *dashed* band is the 95% UCB of $\beta_1(\cdot)$ and the solid horizontal line is $H_0 : \beta_1(\cdot) = 1$. The *dot-dashed* horizontal line is the OLS estimate of fixed β_1 in (5). The GCV-chosen bandwidths are (a) 0.35 (b) 0.36 (c) 0.28 (d) 0.30.



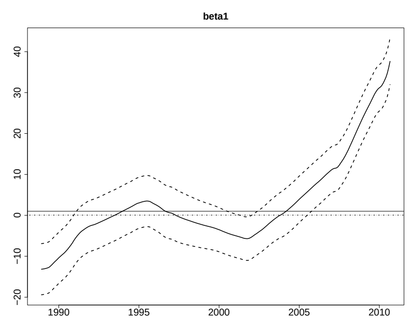
(a) British Pound (GBP)



(b) Japanese Yen (JPY)



(c) Norweigian Krone (NOK)



(d) New Zealand Dollar (NZD)

Figure 2: The *solid* curve is the local-linear regression estimate of $\beta_1(\cdot)$ in (6). The *dashed* band is the 95% UCB of $\beta_1(\cdot)$ and the solid horizontal line is $H_0 : \beta_1(\cdot) = 1$. The *dot-dashed* horizontal line is the OLS estimate of fixed β_1 in (5). The GCV-chosen bandwidths are (a) 0.33 (b) 0.27 (c) 0.27 (d) 0.30.

Table 3: Proportion of $\hat{\beta}_1(t) > 1$

Currency	Proportion
Australian Dollar (AUD)	0.2252
Canadian Dollar (CAD)	0.7557
Swiss Franc (CHF)	0.1832
Danish Krone (DKK)	0.0763
British Pound (GBP)	0.3931
Japanese Yen (JPY)	0.1489
Norweigian Krone (NOK)	0.4046
New Zealand Dollar (NZD)	0.4504

Table 4: Proportion of $\beta_1 = 1$ violating $95\% UCB$

Currency	Proportion
Australian Dollar (AUD)	0.6260
Canadian Dollar (CAD)	0.1565
Swiss Franc (CHF)	0.1489
Danish Krone (DKK)	0.3015
British Pound (GBP)	0.2023
Japanese Yen (JPY)	0.4122
Norweigian Krone (NOK)	0.3969
New Zealand Dollar (NZD)	0.4084

Table 5: Estimates of the coefficients in (10) when lag=1; The numbers inside the brackets are the corresponding *p-values*. The coefficients significant at 5% are in bold.

Currency	US money	Foreign money	US production	Foreign production	T-Bill differential	US money volatility
AUD	−32.88 (0.4802)	**−46.28** (0.0037)	−18.24 (0.3002)	**−165.8** (0.0045)	**0.0279** (3.68×10^{-6})	6.6×10^{-4} (0.6525)
CAD	**126.7** (0.0367)	**84.26** (0.0120)	18.97 (0.4102)	79.67 (0.3153)	**0.0363** (0.0248)	**0.0114** (2.92×10^{-9})
CHF	−10.52 (0.8783)	25.84 (0.3304)	−3.89 (0.8770)	−3.87 (0.8027)	**−0.0887** (0.0003)	**0.0076** (0.0007)
DKK	**−60.75** (0.0395)	**−17.99** (0.0043)	**−54.97** (0.0006)	**4.29** (0.0366)	-6.97×10^{-3} (0.1333)	$\mathbf{-6.63 \times 10^{-3}}$ (4.22×10^{-13})
GBP	**74.46** (0.0077)	−18.52 (0.2822)	3.59 (0.7352)	1.045 (0.7378)	**0.0461** ($<2.00 \times 10^{-16}$)	$\mathbf{-2.69 \times 10^{-3}}$ (0.0019)
JPY	−69.45 (0.1183)	−9.98 (0.2185)	−13.46 (0.4257)	2.258 (0.6056)	**−0.1539** ($<2.00 \times 10^{-16}$)	**0.0053** (0.0003)
NOK	**237.25** (0.0018)	3.5596 (0.9166)	35.35 (0.2998)	−9.999 (0.2228)	**0.0658** (5.57×10^{-6})	**0.0113** (1.93×10^{-6})
NZD	**182.03** (0.0343)	8.9843 (0.8389)	13.08 (0.6910)	−26.64 (0.2997)	0.0247 (0.1228)	**0.0122** (1.56×10^{-5})

Table 6: Estimates of the coefficients in (10) when lag=12; The numbers inside the brackets are the corresponding *p-values*. The coefficients significant at 5% are in bold.

Currency	US money	Foreign money	US production	Foreign production	T-Bill differential	US money volatility
AUD	−53.79 (0.2608)	**−41.62** (0.0095)	−26.54 (0.1359)	**−186.8** (0.0020)	**0.0256** (2.42×10^{-5})	1.91×10^{-3} (0.2094)
CAD	**136.76** (0.0316)	**95.57** (0.0064)	6.0913 (0.7970)	−10.21 (0.9006)	**0.0384** (0.0206)	**0.0117** (7.08×10^{-9})
CHF	33.87 (0.6383)	24.87 (0.3609)	−28.57 (0.2697)	−6.72 (0.6761)	**−0.0662** (0.0083)	**0.0075** (0.0014)
DKK	−43.28 (0.1591)	**−16.74** (0.0100)	**−55.99** (0.0009)	**4.62** (0.0326)	**−0.0111** (0.0201)	$\mathbf{-6.78\times10^{-3}}$ (1.97×10^{-12})
GBP	**107.9** (0.0001)	**−49.95** (0.0124)	4.72 (0.6512)	1.682 (0.5895)	**0.0173** (0.0002)	$\mathbf{-3.10\times10^{-3}}$ (0.0004)
JPY	**−106.9** (0.0352)	−9.116 (0.3136)	−25.62 (0.1761)	2.619 (0.5973)	**−0.0956** (3.90×10^{-8})	$\mathbf{4.17\times10^{-3}}$ (0.0123)
NOK	**274.63** (0.0006)	23.90 (0.4927)	−15.55 (0.6564)	−1.88 (0.8240)	**0.0507** (0.0006)	**0.0112** (7.8×10^{-6})
NZD	**181.94** (0.0351)	10.84 (0.8091)	−15.99 (0.6224)	−26.71 (0.3016)	−0.0158 (0.3164)	**0.0116** (5.6×10^{-5})

Table 7: Wald statistic and p-value for (10) with lag=1

Currency	Wald statistic	p-value
Australian Dollar (AUD)	53.25	< 0.001
Canadian Dollar (CAD)	85.72	< 0.001
Swiss Franc (CHF)	30.19	< 0.001
Danish Krone (DKK)	99.87	< 0.001
British Pound (GBP)	140.23	< 0.001
Japanese Yen (JPY)	111.06	< 0.001
Norweigian Krone (NOK)	51.21	< 0.001
New Zealand Dollar (NZD)	36.62	< 0.001

Table 8: Wald statistic and p-value for (10) with lag=12

Currency	Wald statistic	p-value
Australian Dollar (AUD)	49.11	< 0.001
Canadian Dollar (CAD)	81.58	< 0.001
Swiss Franc (CHF)	28.28	< 0.001
Danish Krone (DKK)	85.32	< 0.001
British Pound (GBP)	51.13	< 0.001
Japanese Yen (JPY)	39.19	< 0.001
Norweigian Krone (NOK)	49.44	< 0.001
New Zealand Dollar (NZD)	45.65	< 0.001

Table 9: Root-mean-squared-error ($RMSE$) for models (12)–(15) under different currencies: For the model averaging, the least squares model averaging ($LSMA$) in Hansen (2007) is used. Each $RMSE$ is divided by the lowest $RMSE$ for the corresponding currency. The best performance is *in bold*.

Currency	Model (12)	Model (13)	Model (14)	Model (15)	LSMA
AUD	1.1489	1.0687	1.1446	1.0020	**1.0000**
CAD	1.1091	1.0278	1.1889	1.0126	**1.0000**
CHF	1.0449	1.0132	1.0447	1.0012	**1.0000**
DKK	1.7982	1.0059	1.9053	**1.0000**	1.0015
GBP	1.1717	1.1285	1.0021	1.1274	**1.0000**
JPY	1.4135	1.0414	1.2281	1.0389	**1.0000**
NOK	1.0430	1.0074	1.0623	1.0032	**1.0000**
NZD	1.0751	1.0304	1.1070	1.0043	**1.0000**