

STUDIES ON TIME SERIES PROPERTIES OF FORWARD DISCOUNT
IN FOREIGN EXCHANGE MARKET

AIDIL RIZAL B. SHAHRIN

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

INSTITUTE OF GRADUATE STUDIES
UNIVERSITY OF MALAYA
KUALA LUMPUR

2015

UNIVERSITI MALAYA ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **AIDIL RIZAL B. SHAHRIN**
Registration/Matric No: **HHB100007**

Name of Degree: **Ph.D**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

**STUDIES ON TIME SERIES PROPERTIES OF FORWARD DISCOUNT IN FOREIGN EXCHANGE
MARKET**

Field of Study: **FINANCIAL ECONOMETRICS**

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature

Date

Subscribed and solemnly declared before,

Witness's Signature
Date

Name:
Designation:

ABSTRACT

Recent literature has suggested that one explanation of the forward bias puzzle is the validity of econometric inference in testing the forward rate unbiasedness hypothesis (FRUH), which results in biased or size distortion. This is due to the highly persistent behaviour of the forward discount. Two models of time series with a highly persistent process are quite successful in explaining the puzzle; long memory and root near unity. However, Choi and Zivot (2007), who focus on long memory, and Sakoulis et al. (2010), who focus on the autoregressive (*AR*) root near unity of the forward discount, report evidence of a less persistent process. These findings are interesting given that the argument in econometrics relies on the highly persistent process of the forward discount. Motivated by these findings, this study extends the discussion of persistency in the forward discount. For that purpose, the monthly forward discount of G7 countries' currencies, which is heavily utilized in earlier studies, is employed as the sample. In the context of long memory, the earlier finding might be spurious due to structural breaks. Based on a semiparametric estimation of the two-step feasible exact local Whittle (FELW) and modified log-periodogram (MLP), we find evidence of long memory before and after structural break adjustment, even though evidence of structural breaks is found in most currencies. Further analysis based on Qu's (2011) and Shimotsu's (2006) tests of spurious long memory reveal no strong evidence in favour of spurious long memory. Overall, the results from structural break adjustment and statistical testing of spurious long memory favour the true long memory process of G7 currencies used in this study. With regard to the mixed finding of stationarity, this study analyses the issue based on a nonlinear framework. To disentangle the issue of nonlinearity and nonstationarity in the forward discount, Caner and Hansen's (2001) model is used. Nonlinearity tests suggest that nonlinearities exist in all G7 countries' currencies. Notably, the forward discount behaves as a unit root in a band and becomes

mean reverting outside the band, which is consistent with the transaction cost argument of nonlinearity creating a ‘band of inaction’. Generally, the forward discount is globally stationary although it is highly persistent since most observations lie inside the dominant unit root regime of the inside band. Thus, previous findings regarding the unit root process for the forward discount must be viewed with caution. In the context of the *AR* root near unity, this study investigates the degree of persistence in the forward discount using the confidence interval approach and Leybourne et al.’s (2007) test of persistency change. A large degree of uncertainty exists in persistency where the upper bound result includes unity in some currencies, which is consistent with the unit root process. Interestingly, the persistency change test reveals that the forward discount undergoes multiple changes in persistence between stationary and nonstationary regimes for most of the currencies. Overall, the findings highlight that the forward discount is not highly persistent throughout the sample period, as previously believed.

ABSTRAK

Kajian terkini telah mencadangkan satu penjelasan mengenai teka-teki bias ke hadapan ialah keabsahan inferens ekonometrik di dalam ujian hipotesis tidak membias kadar ke hadapan (FRUH), yang menghasilkan bias atau herotan saiz. Ini disebabkan oleh kelakuan diskaun ke hadapan yang sangat pesisten. Dua model siri masa dengan proses yang sangat pesisten agak berjaya dalam menjelaskan teka-teki ini; memori panjang dan punca hampir keunitan. Walau bagaimanapun, Choi dan Zivot (2007), yang memberi tumpuan kepada memori panjang, dan Sakoulis et al. (2010) pada punca autoregresif (*AR*) hampir keunitan diskaun ke hadapan, melaporkan bukti proses yang kurang pesisten. Penemuan ini adalah menarik memandangkan hujah di dalam ekonometrik bergantung kepada proses yang sangat pesisten oleh diskaun ke hadapan. Didorong oleh penemuan ini, kajian ini melanjutkan perbincangan mengenai pesisten di dalam diskaun ke hadapan. Untuk tujuan itu, diskaun bulanan ke hadapan matawang-matawang negara-negara G7 yang banyak digunakan di dalam kajian sebelum ini, digunakan sebagai sampel. Di dalam kontek memori panjang, penemuan awal berkemungkinan palsu kerana perubahan struktur. Berdasarkan anggaran dua langkah tempatan tepat semi parametrik Whittle (FELW) dan log-periodogram (MLP) terubahsuai, kita dapati bukti memori panjang sebelum dan selepas pelarasan perubahan struktur, walaupun bukti perubahan struktur didapati di dalam kebanyakan matawang-matawang. Analisis selanjutnya berdasarkan ujian memori panjang palsu Qu (2011) dan Shimotsu (2006) mendedahkan tiada bukti kuat yang memihak kepada memori panjang palsu. Secara keseluruhan, hasil daripada pelarasan perubahan struktur dan ujian statistik memori panjang palsu memihak kepada proses memori panjang benar bagi matawang-matawang G7 yang digunakan di dalam kajian ini. Berhubung dengan dapatan campuran kepegunan, kajian ini menganalisis isu ini berdasarkan rangka kerja tidak linear. Untuk menguraikan isu ketidaklinearan dan ketidakpegunan di dalam diskaun ke hadapan,

model Caner dan Hansen (2001) akan digunakan. Ujian ketidaklinearan mencadangkan bahawa ketidaklinearan wujud di dalam semua matawang-matawang negara G7. Terutamanya, diskaun ke hadapan bertindak sebagai punca keunitan di dalam band dan menjadi purata berbalik di luar band, yang konsisten dengan hujah kos urus niaga bagi ketidaklinearan yang mewujudkan 'band ketiadaan tindakan'. Secara umumnya, diskaun ke hadapan pegun global walaupun ianya sangat pesisten kerana kebanyakan observasi terletak di dalam rejim band punca keunitan yang dominan. Oleh itu, penemuan terdahulu mengenai proses punca keunitan diskaun ke hadapan harus dilihat dengan berhati-hati. Di dalam kontek punca *AR* hampir keunitan, kajian ini menyiasat darjah pesisten di dalam diskaun ke hadapan menggunakan pendekatan selang keyakinan dan ujian perubahan pesisten Leybourne et al. (2007). Ketidakpastian luas wujud di dalam pesisten di mana hasil batas atas termasuk keunitan di dalam beberapa matawang-matawang, yang selaras dengan proses punca keunitan. Menariknya, ujian perubahan pesisten mendedahkan bahawa diskaun ke hadapan mengalami pelbagai perubahan di dalam pesisten diantara regim pegun dan tidak pegun bagi kebanyakan matawang-matawang. Secara keseluruhannya, hasil kajian menyerlahkan bahawa diskaun ke hadapan tidak sangat pesisten sepanjang tempoh sampel, seperti yang dipercayai sebelum ini.

ACKNOWLEDGEMENTS

Alhamdulillah, all praise to Allah s.a.w., this thesis is finally completed. I am grateful to University of Malaya and Ministry of Higher Education of Malaysia for providing me with the scholarship during my postgraduate study. To my supervisors, Dr. Lau Wee-Yeap and Associate Professor Dr. Rubi Ahmad, your assistance and guidance have made the process easier.

Two papers have been accepted for publication by ISI index journal based on the findings from this thesis. Chapter 4 and 5 are accepted for publication in the Journal of Economics and Statistics and Romanian Journal of Economic Forecasting respectively. Useful comments and valuable suggestions from the editors and number of anonymous referees from both journals are highly appreciated. For Chapter 6, it was submitted to Portuguese Economic Journal, an ISI index journal. At the time of this thesis submission, the paper has been returned to the author for revision and re-submission.

Most of the methodologies involves in this thesis-required users' written program. In Chapter 4, the spurious long memory test was conducted using the R statistical software code and MATLAB written by Zhongjun Qu of Boston University and Katsumi Shimotsu of Tokyo University respectively. Whereas for structural break analysis, it is based on the GAUSS code written by Pierre Perron of Boston University and modified log periodogram (MLP) is using the package written in STATA by Christopher F. Baum of Boston College. In Chapter 5, nonlinear unit root test is conduct using GAUSS code written by Bruce Hansen of University of Wisconsin. Forecasting for linear and nonlinear model is using the GAUSS code written by Mehmet Caner of North Carolina State University. For Chapter 6, the grid bootstrap is using the code written by Bruce Hansen of University of Wisconsin, subsampling, written by David E. Rapach of Saint Louis University and the inverse ADF t-statistic written by James Stock

of Harvard University, all written in GAUSS. For persistency test, GAUSS code written by Steve Leybourne of University of Nottingham is used. To all the people involved in writing and making the programs publicly available, your efforts are greatly acknowledged.

Finally, I wish to convey a very special thanks to my family with the unconditional love given regardless whether my research was progressing or stagnating: my father, my mother, brother and sister. To my wife, Siti Salmah Sulaiman, and my lovely daughter, Amirah Nur Sabrina, without your love and affection, the journey of my postgraduate study is far from end. Finally, to Habib Abdullah Salim Al-Jahaf, my spiritual guidance teacher and Ribat Al-Quran's family members who keep on praying for my family success and me today and hereafter, by no means I can repay all of your kindness.

TABLE OF CONTENTS

UNIVERSITI MALAYA ORIGINAL LITERARY WORK DECLARATION	i
ABSTRACT	ii
ABSTRAK	iv
ACKNOWLEDGEMENTS.....	vi
TABLE OF CONTENTS.....	viii
LIST OF TABLES	xi
LIST OF FIGURES	xii
LIST OF ABBREVIATIONS	xiii
Chapter 1: General Introduction.....	1
1.1 Background of The Study	2
1.2 Problem Statements.....	6
1.3 Research Questions	9
1.4 Research Objectives	10
1.5 Contributions.....	11
1.6 Thesis Outline	13
Chapter 2: Review of The Literatures	15
2.1 Efficiency of Foreign Exchange Market.....	15
2.2 Forward Bias Puzzle	20
2.2.1 International Finance	20
2.2.2 Risk Neutrality.....	20
2.2.3 Rational Expectation.....	23
2.2.4 Survey Data	26
2.2.5 Econometric Issues	28
Chapter 3: Methodology and Data	51
3.1 Introduction	51
3.2 Data Sources.....	51
3.3 Forward Rate Unbiasedness Hypothesis (FRUH)	52
3.4 Descriptive Analysis of the G7 currencies.....	54
3.5 Summary	57
Chapter 4: Long Memory of Forward Discount	58
4.1 Introduction	58
4.2 Past Studies on Long Memory	58

4.3 Methodologies.....	62
4.3.1 ARFIMA and Classification of Long Memory Parameter \hat{d}	62
4.3.2 Semiparametric Estimation of Parameter \hat{d}	63
4.3.3 Bai and Perron's (2003) Structural Breaks Model.....	66
4.3.4 Choi and Zivot's (2007) Structural Break Adjustments.....	68
4.3.5 Shimotsu's (2006) 'Simple but Effective' Tests.....	68
4.3.6 Qu's (2011) Spurious Long Memory Test.....	69
4.4 Autocorrelation Function and Periodogram.....	70
4.5 Findings.....	72
4.5.1 Semiparametric Estimates before Break Adjustments.....	72
4.5.2 Results of The Structural Breaks Model.....	75
4.5.3 Semiparametric Estimates after Break Adjustments.....	81
4.5.4 Spurious Long Memory Tests Result.....	83
4.6 Conclusion.....	85
Chapter 5: Nonlinearity in Forward Discount.....	87
5.1 Introduction.....	87
5.2 Stochastic Process of Forward Discount.....	87
5.3 Nonlinearities.....	89
5.4 Methodologies.....	91
5.4.1 Nonlinear Unit Root Test Model.....	91
5.5 Unit Root Test.....	94
5.6 Findings.....	95
5.6.1 Non-Linearity Test.....	95
5.6.2 TAR Unit Root Test.....	96
5.6.3 Out of Sample Forecast.....	100
5.7 Conclusion.....	101
Chapter 6: Autoregressive Root Near Unity Process Of Forward Discount.....	103
6.1 Introduction.....	103
6.2 Root near Unity and Confidence Interval.....	103
6.3 Methodologies.....	105
6.4 Findings.....	112
6.5 Conclusion.....	117
Chapter 7: Conclusion, Implications and Limitations.....	119
7.1 Introduction.....	119
7.2 Summary of The Thesis.....	120

7.3 Implications.....	122
7.4 Limitations	124
REFERENCES.....	127

University of Malaya

LIST OF TABLES

Table 2.1: Selective summary of the empirical literatures on forward bias puzzle	34
Table 3.1: Estimate of the difference regression.....	53
Table 3.2: Descriptive statistics of forward discount.....	54
Table 3.3: U.K.'s monetary policy from 1976 to presents	56
Table 4.1: Estimate of \hat{d}_{MLP} without adjustment for breaks.....	76
Table 4.2: Estimate of \hat{d}_{FELW} without adjustment for breaks.....	77
Table 4.3: Structural break test for forward discount	78
Table 4.4: Result of structural break in forward discount.....	80
Table 4.5: Estimate of \hat{d}_{MLP} with breaks adjustment.....	81
Table 4.6: Estimate of \hat{d}_{FELW} with breaks adjustment.....	82
Table 4.7: Shimotsu's (2006) 'simple but effective' long memory test result	83
Table 4.8: Qu's (2011) spurious long memory test	84
Table 5.1: ADF unit root test	95
Table 5.2: Non-linearity test result.....	96
Table 5.3: TAR unit root test	99
Table 5.4: Out of sample forecasting: Linear $AR(k)$ versus nonlinear TAR model	101
Table 5.5: Ratio.....	101
Table 6.1: Point estimate and 95% confidence intervals for measure of forward discount persistence.....	113
Table 6.2: Multiple persistent change test results	116

LIST OF FIGURES

Figure 3.1: Time series plot of monthly forward discount of G7 currencies.....	55
Figure 4.1: Autocorrelation function (ACF) of G7 currencies	71
Figure 4.2: Periodogram of forward discount.....	72
Figure 4.3: Multiple breaks in mean of forward discount.....	79
Figure 6.1: $I(0)$ and $I(1)$ regimes of forward discount.....	117

University of Malaya

LIST OF ABBREVIATIONS

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
AMEX	American Express Bank
AR	Autoregressive Root
ARFIMA	Autoregressive Fractional Moving Average
ARIMA	Autoregressive Fully Integrated Moving Average
ARMA	Autoregressive Moving Average
BCV	Bootstrap Critical Value
BIC	Bayesian Information Criterion
CAPM	Capital Asset Pricing Model
CH	Caner and Hansen
CIP	Covered Interest Parity
CIR	Cumulative Impulse Response
CPI	Consumer Price Index
CVSR	Conditional Variance of the Spot Rate
DF	Dickey-Fuller
DFT	Discrete Fourier Transform
DGP	Data Generating Process
EH-FX	Expectation Hypothesis Foreign Exchange Market
EH-TS	Expectation Hypothesis Term Structure Interest Rate
EMH	Efficient Market Hypothesis
EML	Exact Maximum Likelihood
EMS	European Monetary System
ESTAR	Exponential Smooth Transition Autoregressive
FCUs	Foreign Currency Units
FELW	Feasible Exact Local Whittle

FIML	Full Information Maximum Likelihood
FRUH	Forward Rate Unbiasedness Hypothesis
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GBP	British Pound
GLS	Generalized Least Square
GMM	Generalized Method of Moments
GPH	Geweke and Porter-Hudak
JP Morgan	John Pierpont Morgan
KPSS	Kwiatkowski et al. test
LIBOR	London Interbank Offered Rate
LS	Least Square
LSTAR	Logistic Smooth Transition Autoregressive
MAIC	Modified Akaike Information Criterion
MLE	Maximum Likelihood Estimation
MLP	Modified Log-Periodogram
MMS	Money Market Services
MPC	Monetary Policy Committee
MSCI	Morgan Stanley Capital International
NA	Not Applicable
OLS	Ordinary Least Square
PPP	Purchasing Power Parity
RE	Rational Expectation
RMSE	Root Mean Squared Error
S & P 500	Standard & Poor's 500
STAR	Smooth Transition Autoregressive
SUR	Seemingly Unrelated Regression
TAR	Threshold Autoregressive
TVC	Time Varying Coefficient

UIP	Uncovered Interest Parity
UK	United Kingdom
US	United States of America
VAR	Vector Autoregressive
VECM	Vector Error Correction Model

University of Malaya

Chapter 1: General Introduction

In the field of economics and finance, any puzzling empirical findings will always result in voluminous studies suggesting solutions based on a variety of arguments. This process starts with a few studies reporting an empirical irregularity, which soon becomes a puzzle as large numbers of studies share the same outcome. Puzzles suggest a problem with the theoretical model, which might be overly simplified or unrealistic in some of its assumptions or, alternatively, an empirical problem might be a purely statistical matter.

In international finance, one irregularity that has dominated the field for almost four decades is the empirical puzzle of forward rate unbiasedness hypothesis (FRUH) testing. Forward rate unbiasedness is a rather simple and intuitive hypothesis suggesting that the current log forward exchange rate f_t should provide an unbiased forecast of the next period's log spot exchange rate s_{t+1} . Stated another way, under unbiasedness, the forward discount $f_t - s_t$ provides an unbiased prediction of the future spot exchange return $s_{t+1} - s_t$. With covered arbitrage holding the forward discount equivalent to the interest differential, a large body of empirical research, which is generally based on regressions of the future spot exchange return on the forward discount popularized by Fama (1984), has consistently rejected unbiasedness. Unfortunately, unbiasedness is not just rejected; the forward discount appears to be a perverse predictor of future spot exchange return, predicting movements in the opposite direction from which they actually occur. This paradoxical finding is also known as the forward bias puzzle.

Several explanations have been offered in the literature as possible causes of the puzzle. Recently, studies have cast doubt on the econometric inference in testing the

hypothesis. On the econometric side, bias or size distortion may arise due to the highly persistent behaviour of the forward discount (i.e. Liu and Maynard, 2005). The highly persistent behaviour of the forward discount as the regressor invalidates the standard statistical inference and may potentially explain the anomalous empirical findings (Maynard, 2006). This has led some studies to focus on the persistency of the forward discount. However, the exact form of persistence in the forward discount has been the subject of some debate (Maynard, 2006). In modelling the persistency of the forward discount, two time series models are used; long memory and root near unity. As Liu and Maynard (2005) note, the appropriate inference procedure to rectify the bias or size distortion depends on the assumptions regarding the form of the persistency model for the forward discount, which means that understanding the persistence of the forward discount is of interest.

1.1 Background of The Study

The forward rate unbiasedness hypothesis relates to the idea of analysing efficiency in the foreign exchange market. Under the efficient market hypothesis (EMH), prices should reflect all information available to market participants that would lead to unfeasibility for market participants to earn excess returns due to speculation. Specifically, the efficient market hypothesis as applied to the foreign exchange market can be divided into two components; (1) market participants whom are being endowed with rational expectation (RE) and (2) risk neutrality of market participants. Hence, foreign exchange market efficiency uses a 'joint hypothesis' to test both components. If the risk neutrality assumption holds, the expected foreign exchange gain from holding one currency must be offset by the opportunity cost in holding funds in this currency

rather than the other. This is the basic principle of a classical topic in international finance, dubbed uncovered interest parity (UIP).

The UIP condition is the building block of most theoretical models in international finance and open macroeconomics models. In foreign exchange trade, the notion of interest parity suggests that the same deposit placed at home or abroad should generate the same return. With a covered arbitrage holding, the UIP condition means that the spread between the forward exchange rate and the concurrent spot exchange rate, or the forward discount, on a given date must equal the difference between the default-free nominal interest rate at home and abroad. Thus, any returns from this interest differential are equalized through these exchange rate movements.

If the forward discount is negatively correlated to the subsequent future spot exchange return, the domestic currency will appreciate rather than depreciate when its nominal interest rate is higher and vice versa. This suggests that the forward discount points in the wrong direction of the ex post movement of the future spot exchange return, which contradicts the UIP condition.

Several approaches have been proposed in the literature to reason out this negative correlation. With capital being perfectly mobile according to the UIP hypothesis, approaches to understanding the puzzle can be categorized into:

- a) International finance (joint hypothesis)
 - i) Risk neutrality
 - Risk premium
 - ii) Rational expectation
 - Peso problem
 - Rational bubbles
 - Rational learning

- b) Survey data
- c) Econometric
 - i) Bias or size distortion in the empirical estimates due to the highly persistent behavior of the forward discount

In the first approach, the puzzles reflect the failure of the joint hypotheses of rational expectations and risk neutrality assumptions to hold. Risk neutrality implies forward exchange rates equal to market expectations of the future spot exchange rates. If risk neutrality for market participants does not hold, it suggests those market participants are risk averse. With risk aversion, the UIP condition is distorted by a risk premium since market participants require a higher return than the interest differential in exchange for the risk of holding the foreign currency. The main issue in this approach is the modelling difficulty associated with the risk premium. Even though progress has been made by modelling the risk premium using the dynamic general equilibrium model, it is still unsuccessful in explaining the magnitude of the failure of unbiasedness (Sarno, 2005).

RE ensures that the expectations for future variables, the exchange rate in this study, incorporate all information available at the time the expectations are formed. Failure of rational expectation leads to three distinctive approaches to explaining the puzzle; the peso problem pioneered by Rogoff (1979) and rational bubbles and learning introduced by Lewis (1989a). However, these approaches are rather unsuccessful in explaining the puzzle. Regarding the peso problem, Evans and Lewis (1995) find that it fails to solve the puzzle. As for rational bubbles, Flood and Hodrick (1990) conclude that studies reporting the existence of rational bubbles in the exchange rate are quite thin. In Sarno et al. (2003), failures in learning in explaining the puzzle are documented.

In testing an aspect of the joint hypothesis (i.e. risk premium), it is assumed that the other aspects hold true (i.e. rational expectation). Due to this joint hypothesis problem, some studies have employed survey data on exchange rate expectation. This type of study is pioneered by Froot and Frankel (1989) and Takagi (1991). In this approach, the results suggest failure of both the joint hypothesis of rational expectation and risk neutrality to hold. This raises another issue on modelling the risk premium and failure of the rational expectation to hold simultaneously. Given that modelling the risk premium itself is already debatable and challenging, research into this approach is not progressing.

Recently, studies on the cause of the puzzle have focused on the econometric issue related to hypothesis testing (i.e. Baillie and Bollerslev, 1994b, 2000; Maynard and Phillips, 2001; Maynard, 2006; Sakoulis and Zivot, 2010). As Baillie and Bollerslev (1994b, 2000), Maynard and Phillips (2001), Liu and Maynard (2005) and Maynard (2006) argue, the potential problem in empirical inference is due to the time series properties of the forward discount. A highly persistent forward discount leads to bias or size distortion in the empirical estimate of the hypothesis.

An appropriate inference procedure is required to rectify the bias or size distortion in the empirical inference. However, this depends on the assumptions made regarding the form of the persistency model of the forward discount (Liu and Maynard, 2005). In modelling the persistency of the forward discount, two time series models are used; long memory and root near unity. Actually, much earlier studies are undecided on whether the forward discount involves a stationary or nonstationary process. Given that nonstationarity of the forward discount is unacceptable theoretically, a common theme in the literature is that the forward discount is borderline nonstationary, which reflects a highly persistent process (Maynard, 2006).

Since high persistency of the forward discount plays a crucial role in econometric issues, few studies have focused on understanding persistence in the forward discount. Interestingly, Choi and Zivot (2007) and Sakoulis et al. (2010) emphasise that the forward discount may not necessarily involve a highly persistent process. These findings may invalidate the previous argument that the forward discount is a highly persistent process. In studies by Choi and Zivot (2007), the focus is on the previous findings that the forward discount is a long memory process. Based on the possibility that long memory might be spurious due to structural breaks, they show that the long memory estimate is less persistent after taking structural breaks into consideration.

In the context of root near unity, Sakoulis et al. (2010) argue that the high persistency of the forward discount based on the autoregressive (*AR*) model is exaggerated due to the presence of structural breaks. By modelling the forward discount as an *AR*(1) process, they find a substantial drop in persistence after allowing for structural breaks based on least squares (LS) estimation. Choi and Zivot's (2007) and Sakoulis et al.'s (2010) findings of a less persistent process for the forward discount motivate this current study. Furthermore, to the best of my knowledge, only these two studies highlight the less persistent process of the forward discount, which makes the study of persistence in the forward discount worth pursuing.

1.2 Problem Statements

Historically, the exact stochastic process of the forward discount remains undecided. Based on unit root testing, previous studies are inconclusive about whether the forward discount involves a stationary or nonstationary process. These mixed findings regarding stationarity cover different currencies and time periods, as discussed in a survey by

Engle (1996). Given these inconclusive findings, more recent studies resort to an intermediate scenario between involving a stationary and nonstationary process with persistence. Some studies describe the forward discount as a long memory process (i.e. Baillie and Bollerslev, 1994b, 2000; Maynard and Phillips, 2001; Choi and Zivot, 2007) while others resort to an autoregressive (AR) process with root near to unity (i.e. Liu and Maynard, 2005; Crowder, 1994, 1995; Evans and Lewis, 1995).

As stated earlier, the highly persistent process of the forward discount might explain the empirical puzzle in previous findings. In the context of a root near unity process, the sampling distribution of Fama's (1994) regression-based test display several unusual features that are not accounted for in the conventional asymptotic theory but have important implications for statistical inference in finite samples (Tauchen, 2001). This suggests potentially serious bias and/or size distortion in the popular approach of regression-based testing of the hypothesis. Alternatively, by modelling the forward discount as a long memory process (fractionally integrated), Fama's (1994) regression-based test constitutes a regression imbalance, where the regressand (future spot exchange return) is a short memory process. Again, the finding of long memory of the forward discount leads to size distortion in the regression-based test (Maynard and Phillips, 2001; Maynard, 2003).

The long memory findings of the forward discount in previous literature might be spurious due to structural breaks or regime switching (i.e. Diebold and Inoue, 2001; Granger and Hyung, 2004). It is important to distinguish between the true and spurious long memory since the statistical inference is quite different between these two processes (see Shao, 2011). Even though the issue of spurious long memory has been reported earlier in the econometric theory literature, only recently have formal statistical

tests of spurious long memory been offered. Two of these tests are by Qu (2011) and Shimotsu (2006).¹

Although Choi and Zivot (2007) address the issue of spurious long memory in the forward discount, their study does not provide formal statistical testing of spurious long memory of a series. Their approach is simply to adjust the forward discount for structural breaks to mimic the spurious long memory process where the long memory parameter is estimated after the adjustment. If no evidence of long memory is found after the adjustment, then the long memory process is spurious. Furthermore, the long memory parameter estimation in Choi and Zivot (2007) relies on the semiparametric method of log-periodogram, which is less efficient than the local Whittle estimator (Robinson, 1995).

In the context of root near to unity, the approach is based on the sum of the autoregressive (*AR*) coefficients. With the sum closer to unity, the process is highly persistent but stationary.² However, the point estimate of the sum, which is obtained by the LS method in previous studies, is biased and quite large (i.e. Ledolter, 2009; Shaman and Stine, 1988). Alternatively, rather than focusing on the point estimate, which is known to be biased, the confidence interval approach provides more flexibility in determining the persistence of the forward discount in the context of the autoregressive (*AR*) process. In this approach, the lower and the upper bounds of the 95% confidence interval will provide the range for the sum of the autoregressive (*AR*) coefficients. Even though constructing the confidence intervals based on the least squares for root near unity process is problematic, Mikusheva (2007) suggests procedures that are valid in constructing the confidence interval for the root near unity process.

¹ Ohanissian et al. (2008) also developed a test of spurious long memory; however, the test assumes that the long memory process is due to aggregation of the short memory process, which does not apply to the forward discount. As discussed in the next chapter, the forward discount is equivalent to the interest differential.

² If the sum is unity, it is called a random walk process which is nonstationary.

Another issue in relation to the root near unity approach is that the process is assumed to hold true for the whole sample period, ignoring the possibility that the forward discount might actually have unit root components in part of the sample period. The econometric inference between the near unit root processes of the regressor against a process with unit roots is quite different. Furthermore, unit root in the forward discount implies statistical imbalances in the regression test as an inescapable conclusion (Liu and Maynard, 2005).

Note that in the above modelling of long memory and root near unity, it is assumed that the data generating process (DGP) of the forward discount is linear. As Enders and Granger (1998) and Taylor et al. (2001) argue, if nonlinearities are prevalent under the alternative of stationarity, a linear unit root test suffers from lack of power, which might explain the mixed findings of stationarity in the earlier studies. One of the arguments that may lead to nonlinearity in the forward discount involves transaction cost based on the model developed by Dumas (1992). The existence of transaction cost will create a 'band of inaction', where inside the band there is no adjustment in deviation from the equilibrium that occurs. However, outside the band, the process becomes mean reverting since the benefits of arbitrage exceed the cost. Thus, the regression-based test that is commonly used is miss-specified due to incorrect functional form.

1.3 Research Questions

In line with the objectives of this study stated in the next section, the research questions are classified based on modelling the persistence of the forward discount:

1) Long memory

- a) Does the forward discount have any structural breaks in the sample period?

- b) Does the forward discount show any evidence of long memory before or after structural break adjustments?
- c) Is the forward discount a true or spurious long memory statistically?

2) Nonlinearity

- a) Is the forward discount a nonlinear process?
- b) Is there a 'band of inaction' due to transaction cost?
- c) Does the nonlinear model perform better than the linear model in forecasting the forward discount?

3) Root near unity

- a) Does the forward discount show evidence of a persistence process based on the confidence interval approach?
- b) Does the root near unity process hold throughout the sample period?

1.4 Research Objectives

The objectives of this study can be classified based on the modelling of persistence in the forward discount:

1) Long memory

- a) To examine the existence of structural breaks in the forward discount
- b) To examine the long memory parameter of the forward discount before and after structural break adjustments using a more efficient semiparametric method

- c) To examine whether the forward discount is a true or spurious long memory process using formal statistical tests

2) Nonlinearity

- a) To examine whether or not the forward discount is a nonlinear process
- b) To investigate the existence of a 'band of inaction' based on transaction cost arguments of nonlinearity
- c) To examine forecasting performance of nonlinear versus linear models of the forward discount

3) Root near unity

- a) To examine the persistence based on the confidence interval approach
- b) To investigate whether or not the root near unity assumption holds true for the whole sample period

1.5 Contributions

The focus of this study is analogous to Choi and Zivot (2007) and Sakoulis et al. (2010), who focus on the persistency of the forward discount. On the econometrics side, for the hypothesis testing to be reliable, an appropriate inference procedure is required to address the highly persistent regressor of the forward discount. Even though modifications of the regression specification, for example, by first-differencing or pre-filtering the forward discount prior to estimation (Roll and Yan, 2000; Newbold et al., 1998) or estimation techniques using robust methods such as sign tests (Maynard, 2003)

have been suggested earlier, they deviate from the original framework of the unbiasedness regression test popularized by Fama (1984).

Far too little attention has been given to the time series property of forward discount. In maintaining the original framework of Fama's (1984) regression-based test model, understanding the persistency of forward discount contributes to the field from the econometric perspective, since the appropriate inference procedure depends on the form of the forward discount's persistence (Liu and Maynard, 2005).

As mentioned in the previous section, study conducted by Choi and Zivot (2007) have addressed the issue of spurious long memory in the forward discount. They still found long memory in forward discount even after removing structural breaks. However, their finding is lack of formal statistical evidence and the long memory parameter estimation used is less efficient. This study will contribute to the issue of long memory in forward discount by applying formal statistical tests for spurious long memory and using more efficient estimator of long memory parameter. Thus, the finding of this study in the issue of long memory will complement Choi and Zivot (2007)'s finding of true long memory process of forward discount.

In study by Sakoulis et al. (2010) whom model the forward discount persistency as root near to unity approach, the point estimate of the sum of the autoregressive (AR) coefficient is biased and quite large as discuss in the previous section. In this study, we apply a much more flexible approach of confidence intervals, avoiding the issue of bias of the point estimate of AR coefficient when the sum is close to unity. Furthermore, we also address the assumption that the root near unity hold true for the whole sample period using much recent econometric method. This is important given that if the root is indeed unity for some of the sample period, this unit root component will produce a downward bias in Fama's (1984) regression (Maynard, 2003).

We also study the possibility that the data generating process (DGP) of forward discount is nonlinear. It should be noted that in the approach of long memory or root near unity in modelling the persistency of forward discount, the DGP is assumed to be linear. Specifically, the DGP assumes that the underlying population of interest of forward discount is linearly generated. The finding of nonlinearity in forward discount may suggest that earlier framework of Fama's (1984) regression is mis-specified due to wrong functional form. The issue of misspecification in regression model has long been noted in the literature (e.g. Deegan Jr, 1976), where the ordinary least square (OLS) estimator will be biased.

1.6 Thesis Outline

This thesis consists of seven chapters. In the first chapter, the introduction covers the background of the study, problem statements, research questions and objectives. Chapter 2 consists of two parts. In the first part, a review of the literature in regards to market efficiency of the foreign exchange market is covered. In the second part, the studies that offer possible explanations of the puzzle from the perspective of a joint hypothesis of international finance, survey data and econometric issues are reviewed.

Chapter 3 contains a description of the data used, the FRUH testing results and descriptive statistics of the forward discount. Chapter 4 focuses on the issue of a true or spurious long memory process in the forward discount. The chapter begins with an introduction and section 4.2 discusses the previous studies on the topic. In section 4.3, the methodologies used in this chapter are discussed. Sections 4.4 and 4.5 will report the findings of this research. Finally, section 4.6 concludes the chapter.

In Chapter 5, the focus is on the issue of nonlinearity in the forward discount. The chapter begins with an introduction, followed by section 5.2 and 5.3, which discusses models that support nonlinearities in the forward discount. Section 5.4 discusses methodologies and Caner and Hansen's (2001) model, followed the results of standard unit root testing in section 5.5. The findings are reported in section 5.6 and include forecasting performance of the nonlinear model used in this chapter against a linear autoregressive (*AR*) process. Section 5.7 concludes the chapter.

In Chapter 6, the focus is on the root near unity process of the forward discount. The chapter begins with an introduction and root near unity, which is followed by the methodology in section 6.3. In this methodology section, Hansen's (1999) grid bootstrap method, Romano and Wolf's (2001) subsampling method and Stock's (1991) inversion of the augmented Dickey-Fuller (*ADF*) *t*-statistic for constructing valid confidence intervals are discussed. Discussion of Leybourne et al.'s (2007) model that determines the existence of a unit root process in a sample is also included in section 6.3. In section 6.4, the findings are reported and the following section concludes the chapter.

Chapter 7 begins with an introduction. In the following section, a summary of the thesis is reported. Implications and limitations of the study are covered in sections 7.3 and 7.4, respectively.

Chapter 2: Review of The Literatures

In this chapter, the discussion begins with foreign exchange market efficiency testing and the puzzle originating from it. The discussion then continues by addressing the literature on the joint hypothesis and survey data as a possible explanation of the puzzle. The problems arising from these approaches are also reviewed. Finally, the econometric issues that arise in testing for the forward rate unbiasedness hypothesis are also reviewed. A summary of selective literatures is also provided in the last section of this chapter.

2.1 Efficiency of Foreign Exchange Market

FRUH testing is based on the principle of market efficiency in the foreign exchange market. Professor Eugene Fama of the Booth Business School, University of Chicago, was the first to develop the efficient market hypothesis in the 1970s. This concept first appears in the finance literature as a survey article entitled 'Efficient Capital Market'.³ The core idea of the theory is that the prices fully reflect all available information in the market; thus, no agents should gain any abnormal profits.⁴

In essence, the theory is associated with the idea of the 'random walk', where all subsequent price changes represent a random departure from previous prices.⁵ The theory of random walk links to the efficient market hypothesis through the efficiency of

³ See Fama (1970).

⁴ The early definition of 'efficient market' is a 'market that adjusts rapidly to information' (Fama et al., 1969). However, the modern definition of 'efficient markets' is 'markets that fully reflect available information', as used in this study (see Fama, 1991).

⁵ See Malkiel (2003).

the market, where there are random sequence of price changes in the market. This in turn results in no informational-based trading that will create abnormal profits.⁶

The theory of the efficient market hypothesis has attracted significant criticism. Many researchers from economics, finance and statistics argue that those stocks prices are partially predictable, where studies based on psychological and behavioural aspects of stock determinations show predictability based on past patterns of stock prices and fundamental valuation.⁷ Nonetheless, researchers have yet to come to a consensus about whether markets, particularly financial markets, are efficient.⁸ As Sewell (2011) notes, the efficient market hypothesis is false, but in spirit it is profoundly true.

In the context of the foreign exchange market, if the market is efficient, the current price of the forward exchange rate for delivery at a specified future date should be an unbiased predictor of the future spot exchange rate. If the forward exchange rate is a biased predictor of the future spot exchange rate, on average, the foreign exchange market will create abnormal profits for market participants. However, in reality, the non arbitrage principle will ensure normal profit for market participants.

Testing for efficiency in the foreign exchange market is based on the uncovered interest parity condition, a theory that links domestic and foreign financial assets. The theory states that regardless of domestic or foreign strategy in equilibrium, both should yield the same amount of returns.⁹

This can be stated as:

$$(1 + i_t) = (1 + i_t^*) \frac{S_{t+1}^e}{S_t} \quad (2.1)$$

⁶ See Lo and MacKinley (2001).

⁷ See Malkiel (2003).

⁸ See Lo (2004).

⁹ With the assumptions of risk neutrality.

where agents are indifferent between investing in the two countries where i_t and i_t^* represent local and foreign nominal interest rates of similar securities, respectively. The variable S stands for the spot exchange rate and the superscript e represents market expectation based on information at time t . After a few manipulations, Eq. (2.1) above can be stated as

$$i_t = i_t^* + \Delta s_{t+1}^e \quad (2.2)$$

This is known as the UIP condition, where the small letter s represents the log of the spot exchange rate.¹⁰

The above activities, however, involve some degree of risk. To eliminate the risk from transactions, the agents may instead use the forward exchange rate, a contract where the parties involved exchange one currency for another at a specified forward rate at a specified maturity date. In this form of transaction, the agents are guaranteed a specified amount of return, which makes the transaction riskless. Thus, in equilibrium:

$$(1 + i_t) = (1 + i_t^*) \frac{F_t}{S_t} \quad (2.3)$$

where the expected spot exchange rate is substituted for the forward exchange rate F . Eq. (2.3) above can be further simplified to become:

$$i_t = i_t^* + (f_t - s_t) \quad (2.4)$$

where f_t is the log of the forward exchange rate and $(f_t - s_t)$ is the forward discount (premium). By combining Eq. (2.2) and Eq. (2.4) above, one can state that:

$$\Delta s_{t+1}^e = f_t - s_t \quad (2.5)$$

This can be simplified further into:

¹⁰ $\Delta s_{t+1}^e = \frac{S_{t+1}^e}{S_t} - 1 \approx \ln\left(\frac{S_{t+1}^e}{S_t}\right)$, a rate of depreciation of domestic currency.

¹¹ $Forward\ discount/premium = \frac{F_t}{S_t} - 1 \approx \ln\left(\frac{F_t}{S_t}\right)$.

$$s_{t+1}^e = f_t \quad (2.6)$$

where in equilibrium the expected spot exchange rate should equal the forward rate. The problem with Eq. (2.5) and Eq. (2.6) is that s_{t+1}^e is unobservable. However, based on rational expectation, the difference between s_{t+1}^e and s_t is the rational expectation forecast error. Converting Eq. (2.5) into regression form results in:

$$\Delta s_{t+1} = \beta_1 + \beta_2(f_t - s_t) + \varepsilon_{t+1} \quad (2.7)$$

where ε_{t+1} is the rational expectation forecast error which means that a *joint hypothesis* of rational expectation and risk neutrality holds. Thus, forward rate unbiasedness jointly implies $\beta_1 = 0$, $\beta_2 = 1$ and $E_t \varepsilon_{t+1} = 0$.

However, earlier testing of foreign exchange market efficiency is based on Eq. (2.6), resulting in a simple regression model of:

$$s_{t+1} = \beta_1 + \beta_2 f_t + \mu_{t+1} \quad (2.8)$$

with the null hypothesis of $\beta_1 = 0$ and $\beta_2 = 1$. Thus, if the findings fail to reject the null, the forward rate is an unbiased predictor of the future spot rate. Regression Eq. (2.8) tests the hypothesis in level versus Eq. (2.7), where the first difference of the series is used, which is also known as the Fama (1994) regression.

Empirical results based on regression Eq. (2.8) are supportive. Testing the efficiency of the foreign exchange market in this level form can be traced back as early as the 1970s. Among those studies are Bilson (1981), Frankel (1978), Stockman (1978), Bilson and Levich (1977) and Kaserman (1973). These studies fail to reject the hypothesis that the forward rate is an unbiased predictor of the future spot rate where they find that the estimated slope coefficient is close to unity.

Estimating regression Eq. (2.8), however, does possess a major problem from a statistical point of view. Both the regressor and regressand in Eq. (2.8) are nonstationary following evidence Messe and Singleton (1982) provide, which results in a spurious regression. Since that study, testing the hypothesis in level as in Eq. (2.8) has been abandoned. Furthermore, as Taylor (1995) note, Eq. (2.8) only holds if $\beta_2 = 1$. Failure to hold results in μ_{t+1} in Eq. (2.8) being replaced with $[(1 - \beta_2)s_t + \mu_{t+1}]$.¹² With nonstationary s_t , this results in a very high variance in the samples. Since the idea of ordinary least squares is to minimize the residual, variance does result in the estimated value of β_2 towards unity regardless of the true value of β_2 .

Due to inherent flaws in estimating regression Eq. (2.8), studies in this area use regression Eq. (2.7) instead. Regression Eq. (2.7) actually involves differencing of the series in Eq. (2.8). This differencing should resolve the issue of spurious regression in Eq. (2.8). The same null hypothesis used in the previous model states that the forward discount is an unbiased predictor of the future spot exchange rate return. Vast numbers of empirical studies conducted using a variety of currencies and time periods are unfavourable to the hypothesis. Even worse is that the estimate of β_2 from Eq. (2.7) is significantly different from zero and generally closer to minus unity than plus unity (Froot and Thaler, 1990). Based on Eq. (2.4), whenever foreign currency is at a discount, we will expect our domestic currency to appreciate to offset the interest differential.¹³ However, what has stunned researchers is that most of the studies have found that if $\beta_2 < 0$ the domestic currency depreciates rather than appreciates to offset the interest differential. This phenomenon is best known as the ‘forward bias puzzle’.

¹² Due to reparameterization of Eq. (2.7) such that $\beta_2 \neq 1$ holds, then $s_{t+1} = \beta_1 + \beta_2 f_2 + [(1 - \beta_2)s_t + \mu_{t+1}]$.

¹³ $f_t - s_t = i_t - i_t^*$.

2.2 Forward Bias Puzzle

As mentioned in the previous section, approaches to solutions to the puzzle can be placed into three categories. In this section, a survey of the literature related to the international finance perspective of joint hypothesis, survey data and econometric issues is presented.

2.2.1 International Finance

Among the various approaches to understanding the puzzle offered in the literature, the context of international finance has received significant attention. Specifically, in this approach, the focus is on the ‘joint hypotheses’ of rational expectation and risk neutrality assumptions. Thus, it may cause risk aversion among the agents, departure from the rational expectation hypothesis or both (Sarno, 2005).

2.2.2 Risk Neutrality

The failure of risk neutrality to hold suggests that the agents are risk averse. The existence of risk-averse agents in the foreign exchange market leads to the uncovered parity condition failing to hold. This is because agents require a higher rate of return than the interest rate differential, as shown in Eq. (2.2). One way to resolve this issue is to include a time-varying risk premium in Eq. (2.2), resulting in:

$$i_t - i_t^* = \Delta s_{t+1}^e + \rho_t \quad (2.9)$$

where ρ_t represents the time-varying risk premium at time t . With this new formulation, arbitrage will ensure that the interest differential should equal the return received from holding the foreign currency plus the risk premium.

Estimation of Eq. (2.7) for FRUH is done using ordinary least squares (OLS), which is problematic if the risk premium is omitted, which results in biased and inconsistent estimates of β_2 . Two conditions can cause the above problems, as Barnhart et al. (1999) mentioned; (i) the forward rate must be a function of an unobservable omitted variable and (ii) the term in the forward rate must be stationary, which can be made so if otherwise. Fama (1984) also argues that the time-varying risk premium is correlated with either the forward discount or interest rate differential. This has encouraged a new strand of studies based on Fama's (1984) arguments.

The main issue in this approach is the modelling difficulty associated with the risk premium. An earlier strand of studies in modelling risk premium depends on simple extensions of the static version of the capital asset pricing model (CAPM).¹⁴ In general, the studies provide evidence that the risk aversion parameter is very large but not significantly different from zero and the restrictions imposed in the model are rejected. This represents the weakness of the static model in modelling risk premium that diverts attention to the dynamic general equilibrium model.

In the context of the dynamic general equilibrium model, the main focus is on the Lucas model, a theoretical model in which representative agents in two countries have identical preferences over two consumption goods but different stochastic endowments of the two goods.¹⁵ However, based on empirical analysis of the Lucas model, the model fails to explain risk premium alone as the source of excess returns in the forward foreign.¹⁶

Another approach in modelling risk premium in the context of general equilibrium models is based on the presence of sticky prices. Engel (1999) highlights

¹⁴ See Frankel (1982), Adler and Dumas (1983), Domowitz and Hakkio (1985), Lewis (1988), Engel and Rodrigues (1989), Giovannini and Jorion (1989) and Engle (1992).

¹⁵ See Hansen and Hodrick (1980), Hodrick (1984), Hodrick and Srivastava (1986), Giovannini and Jorion (1987), Bekaert and Hodrick (1992) and Bekaert (1995).

¹⁶ See Sarno et al. (2003).

the risk that arises due to the covariation of consumption and exchange rates in general equilibrium models with nominal rigidities. Another form of rigidity has also been introduced to induce the risk premium. This type of model incorporates 'limited participation' on the part of agents. Households only enter into arbitrage when the benefits sufficiently exceed the cost (Alvarez et al., 2002). Even though these combinations of multiple costs or rigidities are appealing in explaining the forward bias puzzle, they are still unsuccessful in explaining the magnitude of the failure of unbiasedness.¹⁷

More recently, Chinn and Quayyum (2013), which extended the sample examined in Chinn and Meredith (2004) and Chinn (2006), focusing on the link between monetary policy and the behaviour of UIP deviations. This is an extension of McCallum's (1994) arguments that the puzzle may be the result of simultaneity bias induced by the existence of monetary policy reaction function where the difference in interest is set in order to avoid large current exchange rate movements as well as to smooth interest rate movements. In Chinn and Quayyum (2013), they show the forward bias puzzle is very robust when short-horizon data are used, similar to Chinn and Meredith (2004) and Chinn (2006). However, for long-horizon regression based on longer maturity bond, result in a correct positive sign that are generally closer to unity. The differences between short and long horizon result is explained through macroeconomic model of McCallum (1994) by incorporating a reaction function that result in interest rate to respond to innovation in output and inflation. However, the model assumes that exchange rate forecast is an unbiased predictor of the future spot exchange rate, which contradict the finding based on survey data (e.g. Frankel and Froot, 1987).

¹⁷ For an excellent survey on the failure of risk premium in explaining unbiasedness, see Engel (1996) and Sarno (2005).

Based on a simple two-country general equilibrium model, Baccheeta and van Wincoop (2010) argue that investors make infrequent portfolio decisions, which they will gradually buy the currency as time goes on. This explained the continued appreciation of the currency with higher interest rate raises the expected excess return of the currency. Frankel and Poonawala (2010) found that emerging market currencies are less biased than for advanced countries. This suggests that time-varying risk premium may not result in the bias of previous findings. This is due to the fact that emerging markets are supposed to be more risky than advanced country.

2.2.3 Rational Expectation

The next approach in the literatures focuses on the departure from rational expectation. Rational expectation is an important ingredient in economic decision-making; it is related to forecasting the best approximations of the future of the economy. The goodness of forecasts can be judge by the properties of their forecast errors, which are the differences between a sequence of forecast values of a variable and the actual values. A characteristic of good forecasts is that forecast errors are zero on average. Failure to achieve zero forecast errors on average signals that the forecast is biased. The existence of bias suggests that the forecaster is repeatedly making the same mistake, a mistake that should be eliminated by the forecaster's learning process.

The stability of an economy has profound effects on the ability of the forecasting process. Any instability will affect a forecaster's beliefs about the likelihood of future events. For example, the peso problem may occur when an economy is facing instability. More specifically, the peso problem refers to the situation where agents attach a small probability to a large change in the economic fundamentals, which does

not occur in the sample. In the context of the forward bias puzzle, if markets believe that the exchange rate will fall, until it actually does, the forward exchange rate will remain below the spot value of the exchange rate, where the forward rate embodies the market's expectation.

This perspective has attracted a number of studies in attempts to explain the puzzle.¹⁸ The first to apply this peso problem is Rogoff (1979).¹⁹ He states that a skew in the distribution of the forecast error, which does not exist in the sample, would generate evidence of non-zero excess returns in forward speculation. To achieve this, agents attach a small probability to a large change in the economic fundamentals. This approach, however, has a small sample problem, which fails to explain the overall stylized fact of negative β_2 in Eq. (2.7) above.²⁰ Evans and Lewis (1995) conclude that the peso problem itself cannot solve the forward bias puzzle. They show that the bias introduced by the peso problem is economically significant. Overall, in a survey conducted by Sarno et al. (2003) of the peso problem, a very large number of econometric studies encompassing a very large range of exchange rates and sample periods shows that the estimated uncovered interest parity slope is generally negative and closer to minus unity than plus unity.

Another explanation in the context of departure from rational expectation is the 'rational bubble'. The rational bubble also creates non-zero excess returns even in the existence of risk neutrality in the agents. An explosive path in the exchange rate characterizes the bubble, which progressively takes away from the equilibrium level determined by economic fundamentals. This causes an increase in divergence of the exchange rate from the equilibrium value where investors continue to buy a currency

¹⁸ See Engel (1996) for excellent surveys.

¹⁹ Others include Krasker (1980), Lizondo (1983), Borensztein (1987), Lewis (1988, 1991), Engel and Hamilton (1990), Kaminsky (1993) Evans and Lewis (1995).

²⁰ See Taylor (1995).

even though it is already overvalued with respect to fundamentals, hoping the continuation of the bubble will make it profitable to do so.

In the context of empirical testing of the bubbles, the literature has pursued three approaches: the 'variance bounds' (or excess volatility) tests, specification tests and runs tests. The variance bounds test suggests that evidence of speculative bubbles can be seen through excess volatility of the actual exchange rate relative to the volatility of the exchange rate based on the fundamental solutions. Studies conducted by Huang (1981) and MacDonald and Taylor (1993) argue that excess volatility is indeed present in major exchange rates, which may be due to the presence of rational bubbles. However, Flood and Garber (1980), LeRoy (1984) and Flood and Hodrick (1990) argue that the reliability of the excess volatility tests is questionable. This is because they are conditional on a particular exchange rate model and factors other than rational bubbles can affect the exchange rates.

Another approach is based on the specification test, or specifically Hausman's (1978) test. Meese (1986) fails to reject the no-bubbles hypothesis for the dollar-yen, dollar-mark and dollar-sterling exchange rates from 1973 to 1982. However, the specification test is also conditional on the assumed model of exchange rate determination. As Flood and Garber (1980) argue, the monetary model in general predicts poorly; this results in an omitted variable problem that may potentially bias bubble tests towards rejection of the null hypothesis of no bubbles.

Non-parametric tests have also been used to test for speculative bubbles. Evans (1986) applies a non-parametric test by testing the non-zero median in excess return for the dollar-sterling exchange rate. He finds evidence of speculative bubbles where the null hypothesis of zero median is successfully rejected. Again, the finding of this bubble

is based on a particular model of fundamentals. Flood and Hodrick (1990) conclude that studies finding the existence of rational bubbles in the exchange rate are quite thin.

Learning about regime shifts is another explanation for the anomaly in the context of rational expectation departure. Lewis (1989a, 1989b, 1995) first suggests this idea. In this approach, agents fail to exploit the arbitrage opportunity that exists in the ex post data as they learn their surrounding environments. When rational expectation holds and the agents are assumed to know the underlying distribution of economic disturbance, the forecast errors must be orthogonal to the information set while forming expectations together with zero mean. However, learning about the environment may generate forecast errors having serial correlation with a non-zero mean. The failure of this approach to explain the puzzle is documented in Sarno et al. (2003).

More recently, Aggarwal et al. (2009) argue that forward exchange rates are generally not rational forecasts of future spot rates. After accounting for nonstationarity, non-normality and heteroskedasticity of major currencies (Canada, France, Germany, Japan and UK) for over a quarter of century, their finding deepens the puzzle since those markets are the most liquid with very low trading cost. Study by Chakraborty and Haynes (2008) argue that non-rationality may explain the bias. By using monthly and quarterly data of Australia, Canada, UK and Japan currencies, they found negative covariance between the forecast error and the forward discount. However, the sources of non-rationality is absent in Chakraborty and Haynes's (2008) study.

2.2.4 Survey Data

Based on the literature discussed above, all of the studies focus on one aspect of the joint hypothesis while assuming the other aspect holds true. For instance, studies on risk premium as the cause of the puzzle are based on the assumption that rational

expectation holds. This assumption raises an issue given that each aspect of the joint hypothesis holding true is debatable, as discussed in the previous section. To overcome this issue, some studies have used survey data, where the assumption imposed on the expectation formation of market agents is resolved.

The availability of exchange rate expectations data from Money Market Services (MMS), American Express Bank (AMEX) and the *Economist* have led studies in this context by Frankel and Froot (1987) and Frankel (1988). An advantage of using survey data is that the data make possible individual testing of forecast errors; the rational expectation hypothesis and the hypothesis of a time-varying risk premium are possible. Frankel and Froot (1987) find strong evidence that rejection of the null hypothesis is due to the failure of the rational expectation assumption. They also reject the idea of forward discount prediction errors being due to the time-varying risk premium. This finding is further supported by a study conducted by Dominguez (1986), where he argues that rational expectation is the cause for rejection of the market efficiency hypothesis.

However, a contradictory finding comes from Taylor (1988). By using data provided by Godwin, a UK financial consultancy, he finds that failure of the joint hypothesis is due to the significant risk premium. Liu and Maddala (1992), using a cointegration approach and data from MMS, find mixed evidence of the failure of the joint hypothesis. By using two different set frequencies of data, the result for weekly data show that failure is due to risk premium while for monthly data the failure is caused by both risk premium and rational expectation. This finding contradicts the earlier findings of Frankel and Froot (1987), where both studies use the same source of data from MMS. Thus, by using survey data, the results suggest that failure of the joint hypothesis is due to both the risk premium and the rational expectation hypothesis. This raises the question of how to model risk premium and failure of rational expectation

simultaneously. Given that modelling risk premium itself is already challenging from an international finance perspective, it is still unconfirmed whether the existence of the puzzle is due to the failure of both components of the joint hypothesis to hold.

More recently, Topbas (2014) found that markets are not rational and forward exchange rate is biased predictor of spot exchange rate. By using survey data provided by the Central Bank of Turkey, the finding added to the uncertainty of whether the market is really irrational given that studies based on survey data are still inconclusive.

2.2.5 Econometric Issues

Recently, studies have focused on econometric issues that arise in testing the hypothesis. In testing for forward rate unbiasedness hypothesis, previous studies rely on difference regression²¹

$$\Delta s_{t+1} = \beta_1 + \beta_2(f_t - s_t) + \varepsilon_{t+1} \quad (2.10)$$

rather than regression in level, as stated earlier. Baillie and Bollerslev (1994b, 2000) are the first to highlight the econometric issue in testing the forward rate unbiasedness hypothesis of the above regression. This approach is entirely different in that the focus of previous studies is based on the joint hypothesis of risk neutrality and rational expectation holding.

In regards to the econometric issue, bias or size distortion may arise due to a highly persistent regressor (forward discount) in the above regression (Liu and Maynard, 2005). With a highly persistent regressor, the sampling distribution of the above regression will display several unusual features that are not accounted for in the conventional asymptotic theory but have important implications for statistical inference

²¹ This is Eq. (2.7), reproduced in this section for ease in reference.

in finite samples. Alternatively, another possible interpretation is that the highly persistent behaviour in the forward discount, since not observed in the future spot exchange return, constitutes a regression imbalance (Maynard and Philips, 2001; Maynard, 2003). Again, this regression imbalance leads to size distortion in the regression-based test of Eq. (2.10).

Looking at Eq. (2.10), if the regression is well specified where the disturbance is uncorrelated with the regressor, with $\beta_2 \neq 0$, both the regressand and regressor must be either stationary or nonstationary. If the forward discount is a truly nonstationary process, the exchange rate return must also be nonstationary. However, as Roll and Yan (2000) argue, the noise contributed by ε_t is overwhelming so that only an extremely large sample would detect the presence of unit root in the spot return. The above regressor, however, is substantially less noisy, so its unit root is revealed in a modest sample size. Based on simulation, Roll and Yan (2000) argue that both the spot exchange rate return and the forward discount might have a unit root, but the test statistics are powerless to detect its presence in the spot return because of the noise and modest sample size.

Given that true nonstationarity in the forward discount would be ironic, most studies argue that neither the spot exchange rate return nor the forward discount has unit root. However, a problem arises with near nonstationarity of the forward discount. It is well known that the unit root test has weak power for root near unity; it fails to reject the false hypothesis of a unit root unless the sample is really large. In explaining the puzzle, Tauchen (2001) argues two factors that contribute bias in $\hat{\beta}_2$ are highly persistent in $(f_t - s_t)$ and $cov(f_{t+1} - s_{t+1}, \varepsilon_{t+1}) < 0$. Although $cov(f_t - s_t, \varepsilon_{t+1}) = 0$, there is feedback from the disturbance to the future of regressors making it not

strictly exogenous. With these two factors, the regression estimator of $\hat{\beta}_2$ is upward-biased relative to unity and strongly skewed to the right.

Maynard (2003) argues that the coefficient of the above difference regression has a unit root component in its limit distribution that imparts a bias and skewness to the estimator. This occurs because the differencing procedure used for the regressor variable in Eq. (2.10) fails to induce stationarity on the variable, leaving the forward discount with a nonstationary component. Maynard (2003) argues that for full resolution of the puzzle one must include an explanation of why the forward discount is strongly persistent despite the lack of any observable serial correlation in the exchange rate return. Even though most studies model the forward discount as near the unit root process, this still leads to finite-sample problems. The bias continues to play an important role in finite samples, even when the true coefficient is slightly less than one. Maynard (2003) split the bias of $\hat{\beta}_2$ into four components; two represent nonstationary components of the forward discount and two represent stationary covariation in the forward discount spot rate return. In Maynard's (2003) findings, the nonstationary components of the forward discount exhibit a downward bias on $\hat{\beta}_2$ while the bias from covariation in the forward discount spot rate return forced $\hat{\beta}_2$ to be positive. Similar evidence is also found in Maynard (2003), Roll and Yan (2000), Newbold et al. (1998) and Goodheart et al. (1997) based on the autoregressive root near unity model of the forward discount.

Another possible explanation of the forward bias puzzle from the econometric perspective is due to 'unbalanced regression' where the dependent and independent variables do not share the same time series properties. The arithmetic of the integrated process that imposes a structure on the functional relationship between variables has

implications for the regression equation. In particular, it is the mixing of integrated and stationary variables. This scenario is known as an unbalanced regression.

More precisely;

'An unbalanced regression equation is one [in] which the regressand is not the same order of integration as the regressor or any linear combination of the regressor. A requirement in order to obtain a meaningful estimation with integrated variables is balance in the orders of integration of the variables on the left-hand side and the right-hand side of the regression equation (Maddala and Kim, 1998)

Based on the definition above, regression of Eq. (2.10) requires the change in the future spot exchange rate and the forward discount to be of the same order of integration; otherwise, it is subject to unbalanced regression. Specifically, regression Eq. (2.10) is not well specified if the left-hand side variable (Δs_{t+1}) and the right-hand side ($f_t - s_t$) have different degrees of integration. If the order of the integration of ($f_t - s_t$) is between -0.5 and 0.5 then the forward discount is stationary and the estimate of β_2 in Eq. (2.7) is consistent. Otherwise, if the order of integration is ($f_t - s_t$) > 0.5 , the forward discount is non-stationary, which results in the estimate of β_2 in Eq. (2.10) being inconsistent.

It is well established that the nominal exchange rate behaves as an $I(1)$ process; thus, Δs_{t+1} become an $I(0)$ process. However, the main issue is the degree of integration of ($f_t - s_t$) or the forward discount. In earlier studies, uncertainties exist among the integrated of order zero $I(0)$, or stationarity against the integrated of order one $I(1)$, and nonstationarity in the forward discount. Afterwards, due to this

uncertainty of stationary versus nonstationary, the focus switches to intermediate integration between the $I(0)$ and $I(1)$ process or the long memory process.

In their influential paper, Baillie and Bollerslev (2000) capture the salient features of the above regression with a long memory in the forward discount. They simulate a model of the foreign exchange market with the conditional variance having a long memory inherited from the forward discount. Their finding shows empirical consistency where parameter β_2 in Eq. (2.10) converges slowly to the true value of unity. A similar conclusion is found in Maynard and Phillips (2001); they develop the relevant asymptotic theory based on the findings of Baillie and Bollerslev (2000).

Kellard and Sarantis (2008) combine the statistical issue of long memory in the forward discount with economic reasoning to explain the forward bias puzzle. By employing a rational expectation framework, they show that consumption CAPM implies long memory in the risk premium with the conditional variance of the spot rate helping to explain the analogous behaviour in the forward discount. This finding explains the long-tailed distribution that results in the previous findings of negative β_2 in Eq. (2.10).

In Wang and Wang (2009), they demonstrate that the finding of negative estimate in previous finding does not matter empirically. Based on signal to noise and variance ratio test of Australia, UK, Canada, Euro, Japan, Switzerland and France currencies, the finding does not indicate that the market is inefficient, and the misleading finding has previously been due to statistical estimation. Pippenger (2011) proposed an econometric model that incorporates CIP in Fama's (1984) regression, where he argues that effective arbitrage and CIP are the keys to solve the puzzle. Even though the result is very appealing, several researchers highlight some issues in relation to the proposed econometric model such as lacking economic explanation (Muller,

2011), the error in the econometric model is 'deterministic' (Chang, 2011), having an extreme misspecification and multicollinearity (Baillie, 2011) and the econometric model solution is rather 'abstract' (King, 2011).

University of Malaya

Table 2.1: Selective summary of the empirical literatures on forward bias puzzle

Study	Sample period	Data	Methodologies	Findings/Conclusion
1. Aggarwal et al. (2009)	1973:01-1998:12 (Monthly, 1-, 3-, 6-, 12-month horizon)	<ul style="list-style-type: none"> Forward and spot exchange rate of Canada, France, Germany, Japan and UK. 	<ul style="list-style-type: none"> Nonstationarity of f_t and s_t is examine by using ADF regression Cointegration between f_t and s_t with correction for nonnormality and potential sources of nonstationarity from heteroskedasticity. 	<ul style="list-style-type: none"> Forward rate are not rational forecast even after accounting for nonstationarity, non-normality and heteroskedasticity using parametric and nonparametric test. Deepen the puzzle as the currency market is the most liquid
2. Bacchetta and van Wincoop (2010)	1978:12-2005:12 (Quarterly, 3- month horizon)	<ul style="list-style-type: none"> Spot exchange rate of Germany, UK, Japan, Canada and Switzerland Interest rate of 3-month rates as quoted in the London Euromarkets. 	<ul style="list-style-type: none"> Estimate $\Delta s_{t+1} - (i_t - i_t^*) = \beta_1 + \beta_2(i_t - i_t^*) + \varepsilon_{t+1}$ using SUR Calibrate two-country model which agents make infrequent portfolio decision. 	<ul style="list-style-type: none"> All β_2 are significantly less than zero. Argue the model can explain the puzzle where it matches univariate properties of exchange rate and interest rates (volatility and persistence).
3. Baillie and Bollerslev (1994a)	1/3/1980-28/1/1985 (Daily)	<ul style="list-style-type: none"> Nominal exchange rate of Germany, UK, Japan, Canada, France, Italy and Switzerland 	<ul style="list-style-type: none"> Unit roots Correlogram Cointegration between f_t and s_t Fractional cointegration between f_t and s_t 	<ul style="list-style-type: none"> ACF exhibits very slow decline associated with an $I(1)$ process No cointegration Nominal exchange rate appear to be martingale Exchange rate possesses long memory Influence of to the equilibrium exchange

Study	Sample period	Data	Methodologies	Findings/Conclusion
				rate only vanishes at very long horizon
4. Baillie and Bollerslev (1994b)	1974:01-1991:12 (Monthly)	<ul style="list-style-type: none"> Forward and spot exchange rate of Canada, Germany and UK 	<ul style="list-style-type: none"> Correlogram ARFIMA, estimate through MLE 	<ul style="list-style-type: none"> ACF of forward premium/discount exhibit persistence until 2 years lag ARFIMA(2, d, 0) model. Order of integration are I(0.45), I(0.77) and I(0.55) for Canada, Germany and UK respectively.
5. Baillie and Bollerslev (2000)	1974:01-1991:12 (Monthly)	<ul style="list-style-type: none"> Forward and spot exchange rate of Germany Simulated data 	<ul style="list-style-type: none"> Regression Eq. (2.7) Rolling regression UIP model based on time series properties of forward and spot exchange rate. 	<ul style="list-style-type: none"> $\beta_2 = -2.23$ Substantial variation in the estimated slope coefficient based on rolling regression. The puzzle is due to small sample size and persistent in the forward premium/discount.
6. Bekaert and Hodrick (2001)	1975:01-1997:07 (Monthly)	<ul style="list-style-type: none"> Forward and spot exchange rate of US, UK and Germany. Interest rate of US, UK and Germany 	<ul style="list-style-type: none"> Testing expectations hypothesis in the foreign exchange market (EH-FX), which is equivalent to the unbiasedness hypothesis and expectation hypothesis of the term structure interest rate (EH-TS). $\Delta s_{t+1} = \beta_1 + \beta_2(i_t - i_t^*) + \varepsilon_{t+1}$ Expectation hypothesis using a VAR framework Wald, Lagrange and distance metric test 	<ul style="list-style-type: none"> Wald test over-reject the null Lagrange slightly under-reject the null Distance metric over-reject the null Severe size distortion in Wald test, Lagrange multiplier test perform best Unlikely Expectation Hypothesis is true

Study	Sample period	Data	Methodologies	Findings/Conclusion
7. Chakraborty and Haynes (2008)	1998:12-2005:09 (Monthly) 1988:Q4-2005:Q3 (Quarterly)	<ul style="list-style-type: none"> Forward and spot exchange rate of Australia, Canada, UK and Japan. 	<ul style="list-style-type: none"> Cointegration of spot and forward rate Regression Eq. (2.7) Regression Eq. (2.8) General model 	<p>because the requirement of constant risk premium</p> <ul style="list-style-type: none"> The difference between regression Eq. (2.8), the 'level' regression, and regression Eq. (2.7) due to the existence of bias together with the non-stationarity of spot and forward rate. Non-rationality may explain the bias that generates the puzzle.
8. Chinn (2006)	1980:Q1-2000:Q4 (Quarterly, long horizon 3-,6- and 12-month)	<ul style="list-style-type: none"> Spot exchange rate of German, Japan, UK, France, Italy and Canada. Spot exchange rate of emerging countries. Euro currency yields 	<ul style="list-style-type: none"> Long-horizon interest parity regression 	<ul style="list-style-type: none"> Evidence against uncovered interest parity in the current floating era is not as great as is commonly thought. However, for major currencies, short-term interest differential remains a biased predictor of spot return.
9. Chinn and Meredith (2004)	1980:Q1-2000:Q4 (Quarterly)	<ul style="list-style-type: none"> Spot exchange rate of German, Japan, UK, France, Italy and Canada. Interest rate on longer maturity bonds 	<ul style="list-style-type: none"> Using regression $\Delta s_{t+1} = \beta_1 + \beta_2(i_t - i_t^*) + \varepsilon_{t+1}$ Macroeconomic models 	<ul style="list-style-type: none"> Found differences between the tests of UIP using short against long horizon data. Non-standard explanation of the puzzle (common explanations of the puzzle include risk premium, expectational errors or peso problems)

Study	Sample period	Data	Methodologies	Findings/Conclusion
				<ul style="list-style-type: none"> • Suggest the temporary disturbance to the UIP relationship as the cause of the puzzle. • However, the model fails to explain the puzzle found using survey data
10. Chinn and Quayyum (2013)	1980:Q1-2011:Q4 (Quarterly)	<ul style="list-style-type: none"> • Spot exchange rate of Canada, UK, Japan and Switzerland. • Interest rate on longer maturity bonds 	<ul style="list-style-type: none"> • Using regression $\Delta s_{t+1} = \beta_1 + \beta_2(i_t - i_t^*) + \varepsilon_{t+1}$ • Short horizon vs long horizon 	<ul style="list-style-type: none"> • Hold better at longer horizon than at short • Weaker result found in Japan as attributed to low interest rate with zero interest bound which highlights a changing behaviour of the exchange risk premium..
11. Choi and Zivot (2007)	1976:01-1999:01 (Monthly)	<ul style="list-style-type: none"> • Forward and spot exchange rate of Germany, France, Italy, Canada and UK. 	<ul style="list-style-type: none"> • Unit root test of ADF • ARFIMA estimate through semiparametric (MLP) • Structural break • Break adjustment 	<ul style="list-style-type: none"> • Fail to reject unit root in forward discount of Germany, Canada and UK • All forward discount integrated of the order $I(0.54) \leq I(d) \leq I(0.87)$, a nonstationary long memory process. • All forward discount experience multiple breaks • After break adjustment, all forward discount integrated of the order $I(0.23) \leq I(d) \leq$

Study	Sample period	Data	Methodologies	Findings/Conclusion
12. Clarida et al. (2003)	7/1/1979-31/12/1998 (Weekly, 4-, 13-, 26- and 52-week forward)	<ul style="list-style-type: none"> Forward and spot exchange rate of France, Germany, Japan and UK. 	<ul style="list-style-type: none"> Unit root test of ADF Cointegration between s_t and $f_t^4, f_t^{13}, f_t^{26}$ and f_t^{52} Regime-switching VECM 	<p>$I(0.46)$, a stationary long memory process.</p> <ul style="list-style-type: none"> Fail to reject the null of unit root Strongly reject hypothesis of 3 independent cointegrating vectors against alternative of four. Exchange rate dynamic display nonlinearities
13. Crowder (1994)	1974:01-1991:12 (Monthly)	<ul style="list-style-type: none"> Forward and spot exchange rate of UK, Germany and Canada 	<ul style="list-style-type: none"> ADF unit root test on nominal and forward exchange rate Cointegration between spot rate of all currencies ADF unit root test on forward premium/discount 	<ul style="list-style-type: none"> All nominal and forward exchange rate possess a unit root Null of no more than 1 and no more than 2 cointegration vectors cannot be rejected UK forward premium/discount is stationary, but not Germany and Canada Common stochastic trend is not an instrument for forward risk premium Foreign exchange market efficiency may hold if forward premium/discount is a long memory stationary process.

Study	Sample period	Data	Methodologies	Findings/Conclusion
14. Crowder (1995)	27/9/1982-30/3/1990 (Monthly)	<ul style="list-style-type: none"> Eurocurrency deposit rate of US, UK, Germany, Japan, Canada and France Forward and spot exchange rate of US, UK, Germany, Japan, Canada and France 	<ul style="list-style-type: none"> ADF unit root test on interest differential and forward premium/discount VECM 	<ul style="list-style-type: none"> Interest differentials contain unit roots All forward premium/discount reject unit root except UK Forward premium/discount and interest differential are cointegrated Market are efficient except for Germany-US
15. Dominguez (1986)	1983-1985 (Weekly mix)	<ul style="list-style-type: none"> 30 respondents of exchange rate forecaster for exchange rate expectation Spot exchange rate of UK, Germany, Switzerland and Japan . 	<ul style="list-style-type: none"> $\Delta s_{t+1} = \beta_1 + \beta_2(E_t s_{t+1} - s_t) + \varepsilon_{t+1}$ where $E_t s_{t+1}$ is the expectation of future spot rate. 	<ul style="list-style-type: none"> Reject the rational expectation hypothesis
16. Evans and Lewis (1995)	1975:01-1989:12 (Monthly)	<ul style="list-style-type: none"> Forward and spot exchange rate of Germany, UK and Japan 	<ul style="list-style-type: none"> Unit root test Switching model Cointegration 	<ul style="list-style-type: none"> All forward discount are nonstationary Anomalous behaviour can be explained by rational expectation Peso problem explains the apparent permanent shock to the risk premium which can induce bias in regression Eq. (2.7)
17. Fama (1984)	1973:08 – 1982:12 (Monthly)	<ul style="list-style-type: none"> Forward and spot exchange rate of Belgium, Canada, France, Italy, Japan, Netherlands, Switzerland, 	<ul style="list-style-type: none"> Regression equation of $s_{t+1} - s_t = \alpha_1 + \beta_1(f_t - s_t) + \varepsilon_{1,t+1}$ $f_t - s_{t+1} = \alpha_2 + \beta_2(f_t - s_t) + \varepsilon_{2,t+1}$ 	<ul style="list-style-type: none"> All β_1 are significantly negative Variation in forward rate is due to variation in

Study	Sample period	Data	Methodologies	Findings/Conclusion
		UK and West Germany	<ul style="list-style-type: none"> • SUR estimate • Sub-period analysis 	<ul style="list-style-type: none"> • premium of forward discount or expected change in the spot rate. • Risk premium and expected future spot rate are negatively correlated
18. Frankel and Froot (1987)	1976-1979 (for Amex) 1981-1985	<ul style="list-style-type: none"> • Survey data from MMS, Economist and Amex 	<ul style="list-style-type: none"> • Extrapolative, adaptive and regressive model of expectation. 	<ul style="list-style-type: none"> • Static and bandwagon expectation are rejected • Significant bias is found in the process governing bias.
19. Frankel and Poonawala (2010)	1996:12-2004:04 (Monthly)	<ul style="list-style-type: none"> • Forward and spot exchange rate of industrialized and emerging countries. A total of 35 currencies. 	<ul style="list-style-type: none"> • Regression Eq. (2.7) for each country • Pooled regression • SUR estimate 	<ul style="list-style-type: none"> • Strong forward rate bias (negative β_2) evidence for industrialized countries • Emerging countries are less negative than industrialized countries • Pooled analysis of emerging countries is $\beta_2 = -0.028$, for advanced $\beta_2 = -2.023$ • Argue that time-varying risk premium may not be the explanation of the bias.
20. Froot and Frankel (1989)	1976 – 1985 (Monthly mix)	<ul style="list-style-type: none"> • Survey data on exchange rate expectation from MMS and AMEX 	<ul style="list-style-type: none"> • Regression Eq. (2.7) • Decompose the bias into portion attributable to risk premium and expectational errors. 	<ul style="list-style-type: none"> • All β_2 are significantly less than zero. • None of the bias reflect risk premium
21. Giovannini and Jorion	5/7/1974 – 19/12/1986	<ul style="list-style-type: none"> • Spot exchange rate of US, 	<ul style="list-style-type: none"> • Model with constant variance 	<ul style="list-style-type: none"> • Estimated conditional

Study	Sample period	Data	Methodologies	Findings/Conclusion
(1989)	(Weekly)	<ul style="list-style-type: none"> German, UK and Switzerland. Stock of government debt 	<ul style="list-style-type: none"> Model with time varying covariance 	<ul style="list-style-type: none"> variances cannot explain the observed time variation of risk premium.
22. Kellard and Sarantis (2008)	1991:01-2005:10 (Monthly)	<ul style="list-style-type: none"> Spot exchange rate of US Dollar/Sterling, Yen/US Dollar, Deutschmark/US Dollar, Deutschmark/Sterling, Deutschmark/Yen, Euro/Yen, Euro/Sterling and Euro/US Dollar. 1-month interest rate differentials Square root of 'traded' implied volatilities as proxy of conditional variance of the spot rate (CVSR) 	<ul style="list-style-type: none"> ARFIMA estimate through GPH Bai and Perron (2003) structural breaks Choi and Zivot (2007) demean procedure Cointegration 	<ul style="list-style-type: none"> GPH estimate for forward premium/discount are $I(0.85) \leq I(d) \leq I(0.99)$, CVSR $I(0.65) \leq I(d) \leq I(0.94)$ Forward premium/discount, 2-5 breaks are found while CSVSR, 1-3 breaks After demean, long memory estimate show little difference. Forward premium/discount and CVSR are characterised as both long memory and structural breaks. Based on the demean series, forward premium/discount and CVSR are fractionally cointegrated.
23. Killian and Taylor (2003)	1973:Q1-1998:Q4 (Quarterly)	<ul style="list-style-type: none"> Nominal exchange rate of Canada, France, Germany, Italy, Japan, Switzerland and UK. US and foreign CPI 	<ul style="list-style-type: none"> ESTAR Long horizon regression test Bootstrap test of the random walk hypothesis 	<ul style="list-style-type: none"> Found nonlinearity Found strong evidence of predictability at horizons of 2 to 3 years, but not at shorter horizons.

Study	Sample period	Data	Methodologies	Findings/Conclusion
24. Lewis (1989a)	1973:07 – 1984:06 (Monthly)	<ul style="list-style-type: none"> • Spot exchange rate of German, Japan and UK • Interest rate 	<ul style="list-style-type: none"> • Learning forecast error model • Monetary model 	<ul style="list-style-type: none"> • The finding helps explain the PPP puzzle • Change in parameter of the money market affects the behaviour of exchange-rate forecast error while the market is learning. • Forecast error would be on average negative.
25. Liu and Maynard (2005)	1979:10-1986:11 (Monthly for Japan and Australia) 1973:7-2000:3 (Monthly for Canada, France, Germany and UK)	<ul style="list-style-type: none"> • Spot exchange rate of Japan, Australia, Canada, France, Germany and UK. • Forward rate calculated using the LIBOR 1-month interest differential 	<ul style="list-style-type: none"> • Confidence interval by inverting ADF test by Stock (1991) • Sup-bound method • Bonferroni method • Scheffe method 	<ul style="list-style-type: none"> • Most confidence interval including 1, with exception to UK. This uncertainty regarding the degree of persistence suggests possible size distortion in testing $\beta_2 = 1$ of Eq. (2.7). • It explains potential over-rejection when using standard OLS in regression Eq. (2.7). • After accounting for possible near unit root behaviour of forward premium/discount, evidence against unbiasedness is found for bot full and sub-samples when using Bonferroni and Sup-bound method. Scheffe method also fails to reject for Australia,

Study	Sample period	Data	Methodologies	Findings/Conclusion
				Germany, Japan and Canada. <ul style="list-style-type: none"> Size distortion may not provide the sole explanation of the puzzle.
26. Mankiw and Shapiro (1986)	NA	<ul style="list-style-type: none"> Simulated data of sample size $N = 50$ and $N = 200$. Typical size found in applied macroeconomic research. 	<ul style="list-style-type: none"> Study standard test model of regression of the variable Y_t on lagged information (as Eq. (2.7)) of $Y_t = \Phi_0 + \Phi_1 X_{t-1} + v_t$ and evaluate the model when testing the null of $H_0 = \Phi_1 = 0$ $X_t = \theta X_{t-1} + \epsilon_t$, under null, X_{t-1} and v_t are uncorrelated. However, does not preclude a contemporaneous correlation between ϵ_t and v_t. Thus, Gauss-Markov theorem does not apply. The justification for OLS and subsequent hypothesis testing relies on asymptotic distribution theory. 	<ul style="list-style-type: none"> Using asymptotic distribution will reject a true null hypothesis more than 5% of the time. If θ is close to unity (this the case of forward discount), a valid rejection at a 5% significance level require much more conservative critical value.
27. Maynard (2003)	1986:11-1998:03 (Daily)	<ul style="list-style-type: none"> Spot exchange rate of Australia, Canada, France, Germany, Japan and UK. Forward rate calculated using the London Eurocurrency 1-month interest differential 	<ul style="list-style-type: none"> Unit roots for both spot s_t and forward rate f_t Cointegration between s_{t+1} and f_t Split the bias in regression Eq. (2.7) into: <ul style="list-style-type: none"> b_1: Effects of Dickey-Fuller distribution. b_2: one-sided long run covariance b_3: short-run covariance b_4: covariance between the Δs_{t+1} and stationary 	<ul style="list-style-type: none"> Confirm nonstationarity in level for both spot and forward rate Null hypothesis of no cointegration is rejected at 5% level for all currencies. Above results imply unit root behaviour in the forward discount Result from $b_1 + b_2$ are negative. Nonstationary

Study	Sample period	Data	Methodologies	Findings/Conclusion
			<p>component of $f_t - s_t$, the forward discount</p> <ul style="list-style-type: none"> b_1 and b_2, unit-root-type biases arising from the nonstationary component of forward discount. b_3 biases reflect covariation between stationary component of forward discount and spot return. 	<p>behaviour of forward discount exhibits a downward bias on β_2.</p> <ul style="list-style-type: none"> b_3 is negative, but fairly small in magnitude. b_4 is positive This capability of explaining $\hat{\beta}_2 < 1$, but unlikely to generate $\hat{\beta}_2 < 0$. The limit theory can only partially account for the negative empirical estimates.
28. Maynard (2006)	<p>1979:09-1986:10 (Monthly data for Japan and Australia)</p> <p>1973:06-2000:03 (Monthly data for Canada, France, Germany and UK)</p>	<ul style="list-style-type: none"> Spot exchange rate of Japan, Australia, Canada, France, Germany and UK. Forward rate calculated using the LIBOR 1-month interest differential 	<ul style="list-style-type: none"> Estimate the equation of : $s_{t+1} - s_t = \alpha + \beta(f_t - s_t) + e_{t+1}$ $s_{t+1} - f_t = \alpha + (\beta - 1)(f_t - s_t) + e_{t+1}$ Unit root and persistency of $(f_t - s_t)$ and $(s_{t+1} - f_t)$ Sign tests, test for sign predictability $(s_{t+1} - f_t)$ using $(f_t - s_t)$, which imply a violation of unbiasedness with null no predictive power. Covariance estimate and test. Conditional test 	<ul style="list-style-type: none"> The results for the equation show strong evidence against unbiasedness. Unit root rejected for $(s_{t+1} - f_t)$, only 2 rejections for $(f_t - s_t)$. MLP estimate for $(s_{t+1} - f_t)$ is relatively small, $(f_t - s_t)$ is large. Size correction is available, but requires the researcher to take stands on the nature of persistence (long memory, root near unity or structural breaks). Sign test indicates negative sign, confirm the traditional finding of

Study	Sample period	Data	Methodologies	Findings/Conclusion
29. Maynard and Phillips (2001)	1986:11-1998:03 (Daily)	<ul style="list-style-type: none"> Forward and spot exchange rate of Australia, Canada, France, Germany, Japan and UK. 	<ul style="list-style-type: none"> ARFIMA, estimate through MLP and MLE. Statistical issue of regression Eq. (2.7) 	<ul style="list-style-type: none"> $\hat{\beta}_2 < 1$. Covariance tests are qualitatively similar to those from the sign test. Conditional test delivers strongest rejections shows evidence against unbiasedness. Forward discount can be model as a non-stationary fractionally integrated process with $0.5 < d < 1$. If $(f_t - s_t)$ is random walk and $s_{t+1} - s_t$ is linear trend, the standard t statistic diverges. The non-stationary long memory of forward discount suggests a statistical imbalance in regression Eq. (2.7). Long memory of forward discount provides itself a rejection of forward rate unbiasedness. The presence of long memory in forward discount, unnecessary to conduct further tests of unbiasedness. Regression Eq. (2.7)

Study	Sample period	Data	Methodologies	Findings/Conclusion
				<p>must be adjusted to account for the stochastic trend in forward discount.</p> <ul style="list-style-type: none"> • Average value of β_2 is -4 • The unbiasedness of forward rates as predictor of future spot exchange rate (Eq. (2.8)) is not identical to Eq. (2.7). So unbiasedness rejections are not conclusive for UIP. • Simultaneity bias
30. McCallum (1994)	1978:01 – 1990:07 (Monthly)	<ul style="list-style-type: none"> • Forward and spot exchange rate of Japan, German and UK 	<ul style="list-style-type: none"> • Regression Eq. (2.7) and (2.8) • ARMA • Monetary policy model 	<ul style="list-style-type: none"> • It suggests that s_t and f_t do not have stable univariate AR representations, even after removing a linear trend. • It suggests that $(s_{t+1} - f_t)$ have stable univariate AR representations for Canada and Germany (no unit root). However, for Switzerland the result are consistent with presence of a single unit root. • Natural logarithms of exchange rates are well
31. Meese and Singleton (1982)	7/1/1976-24/6/1981 (Weekly)	<ul style="list-style-type: none"> • Forward and spot exchange rate of Switzerland, Germany and Canada. 	<ul style="list-style-type: none"> • Test of unit roots in the AR representation of s_t and f_t. • Test of unit roots in the AR representation of $(s_{t+1} - f_t)$. 	<ul style="list-style-type: none"> • It suggests that $(s_{t+1} - f_t)$ have stable univariate AR representations for Canada and Germany (no unit root). However, for Switzerland the result are consistent with presence of a single unit root. • Natural logarithms of exchange rates are well

Study	Sample period	Data	Methodologies	Findings/Conclusion
				<p>approximated by random walks.</p> <ul style="list-style-type: none"> Regression Eq. (2.8) may be inappropriate, since the asymptotic distribution theory employed may not be valid.
32. Muller (2011)	1983:11-1993:09 (Weekly, this sample is taken from Pippenger (2011) 2nd sub-sample)	<ul style="list-style-type: none"> Forward and spot exchange rate of UK. 	<ul style="list-style-type: none"> Focus on solution of the puzzle by Pippenger (2011). Pippenger (2011) suggest an econometric model of $\Delta s_{t+1} = \lambda_0 + \lambda_1(f_t - s_t) + \lambda_2(f_{t+1} - f_t) - \lambda_3(i_{t+1} - i_{t+1}^*) \pm e_{t+1}$ 	<ul style="list-style-type: none"> Pippenger (2011) econometric exercise took stationarity of $(f_t - s_t)$ or forward discount for granted. Disagree with Pippenger (2011) claims that omitted variable bias is the solution to the forward bias puzzle. Pippenger (2011) lack of economic model, need to look for fundamentally better models of foreign exchange rate determination.
33. Newbold et al. (1998)	1984:05-1995:10 (Monthly)	<ul style="list-style-type: none"> Forward and spot exchange rate of Germany, Japan and UK 	<ul style="list-style-type: none"> AR process of $(s_{t+1} - s_t)$ and $(f_t - s_t)$ Cointegration 	<ul style="list-style-type: none"> Time series properties of $(s_{t+1} - s_t)$ and $(f_t - s_t)$ are quite different. The former exhibits only mild autocorrelation, the latter appears to be generated by a process with a very large, possible unit,

Study	Sample period	Data	Methodologies	Findings/Conclusion
				<p>autoregressive root.</p> <ul style="list-style-type: none"> • Forward discount is only a very small component of excess return. • Stationarity of $(f_t - s_t)$ permits the use of regression Eq. (2.7) • Paradox disappears when one allows for autocorrelation in the series of excess return.
34. Pippenger (2011)	<p>5/1/1961-29/6/1973 (Weekly, for Canada) 2/1/1974-30/9/1993 (Weekly, for UK)</p>	<ul style="list-style-type: none"> • Forward and spot exchange rate of Canada and UK • 1-month euro rates 	<ul style="list-style-type: none"> • Suggest econometric model as solution of the puzzle: $\Delta s_{t+1} = \lambda_0 + \lambda_1(f_t - s_t) + \lambda_2(f_{t+1} - f_t) - \lambda_3(i_{t+1} - i_{t+1}^*) \pm e_{t+1}$ • Bias due to omitted variables of $(f_{t+1} - f_t)$ and $(i_{t+1} - i_{t+1}^*)$. 	<ul style="list-style-type: none"> • Regression Eq. (2.7) is missing 2 variables that CIP implies should be included. • Those missing variables explain the downward bias in the puzzle. • The variance of changes in exchange rates is 100-200 times larger than the variance in forward discount.
35. Sakoulis et al. (2010)	<p>1976:01-1998:12 (Monthly)</p>	<ul style="list-style-type: none"> • Forward and spot exchange rate of Germany, France, Italy, Canada, UK and Japan. 	<ul style="list-style-type: none"> • The AR(1) model of forward discount $f_t - s_t = c + \Phi(f_{t-1} - s_{t-1}) + v_t$ • Bai and Perron (2003) break model • Multiple structural break model of $f_t - s_t = c_j + \Phi(f_{t-1} - s_{t-1}) + u_t,$ $t = T_{j-1} + 1, \dots, T_j$ • Monte Carlo experiments 	<ul style="list-style-type: none"> • Estimate of Φ from OLS are $0.70 \leq \Phi \leq 0.94$. • All currencies, except Japan, experience multiple breaks in the AR(1) representative of forward discount. • All the point estimate of the Φ coefficients across countries has dropped

Study	Sample period	Data	Methodologies	Findings/Conclusion
				<ul style="list-style-type: none"> significantly. High persistence due to structural breaks. This is a plausible alternative to the fractionally cointegrated.
36. Sarno et al. (2006)	4/1/1985-31/12/2002 (Weekly, 4- and 13-week)	<ul style="list-style-type: none"> Forward and spot exchange rate of Japan, UK, Germany, Euro and Switzerland. 	<ul style="list-style-type: none"> Regression Eq. (2.7) Nonlinear Fama regression using STAR 	<ul style="list-style-type: none"> All β_2 are significantly less than zero. Major bilateral dollar exchange is linked nonlinearly to forward premia. If the true process of UIP deviations were nonlinear, it may explain the puzzle.
37. Sercu et al. (2008)	1/6/1985-1/4/1998 (Daily and weekly)	<ul style="list-style-type: none"> Forward and spot exchange rate of Belgium, Denmark, France, Netherlands, Switzerland, Italy, UK, Japan, Canada and US. 	<ul style="list-style-type: none"> Cumby-Obstfeld-Fama test of UIP GARCH FIML and GMM 	<ul style="list-style-type: none"> Zero coefficient for intra-EMS rate Negative coefficient for extra-EMS rate Reject the hypothesis of either a risk premium or peso issue as sufficient explanations.
38. Tauchen (2001)	1978:07-1981:12 (Weekly)	<ul style="list-style-type: none"> Forward and spot exchange rate of UK. 	<ul style="list-style-type: none"> Small sample properties of β_2 in regression Eq. (2.7) 	<ul style="list-style-type: none"> Under expectation theory, the sampling distribution of the regression estimator of β_2 of this coefficient is upward-biased relative to unity and strongly skewed to the right. Thus the biased in a

Study	Sample period	Data	Methodologies	Findings/Conclusion
				<ul style="list-style-type: none"> direction opposite to previously observed. The evidence against the hypothesis of unbiased forward rates is much stronger than previously believed.
39. Topbas (2014)	2002:01-2006:04 (Monthly survey data)	<ul style="list-style-type: none"> Turkey's expectation regarding US against Turkey spot exchange rate 	<ul style="list-style-type: none"> $\Delta s_{t+1} = \beta_1 + \beta_2(E_t s_{t+1} - s_t) + \varepsilon_{t+1}$ where $E_t s_{t+1}$ is the expectation 	<ul style="list-style-type: none"> Mean of $(E_t s_{t+1} - s_t)$ is higher than mean of $(f_t - s_t)$. Reject regression Eq. (2.7). Forward discount contains additional information for exchange rate forecasts.
40. Wang and Wang (2009)	1985:01-2004:12 (Monthly)	<ul style="list-style-type: none"> Forward and spot exchange rate of Australia, UK, Canada, Euro, Japan, Switzerland and France. 	<ul style="list-style-type: none"> Regression Eq. (2.7) Signal to noise ratio Variance ratio 	<ul style="list-style-type: none"> All β_2 are significantly less than zero. Do not indicate market are inefficient Demonstrate that absolutely unbiased predictor is irrelevant empirically.

Chapter 3: Methodology and Data

3.1 Introduction

In this chapter, the discussion begins with the data used in this study. The discussion then continues by testing the Fama's (1984) regression followed by the descriptive statistics of the forward discount. Finally, the chapter concludes with the summary.

3.2 Data Sources

In this study, all the data sets are obtained from Datastream where originally sourced by WM Company/Reuters. There are two types of exchange rate data involved; the 1-month forward exchange rate and spot exchange rate of G7 countries' currencies. Each exchange rate of G7 countries' currencies quoted as foreign currency units (FCUs) per British pound (GBP). Accordingly, the six currencies involved in this study are Canadian dollar, French Franc, German mark, Italian lira, Japanese yen and U.S. dollar.

Formerly, the data consists of daily observations on all weekdays of dealer quotes of bid and ask spot exchange rates and 1-month forward exchange rates. The daily data then converted into non-overlapping monthly observations of the last weekday of each month. Based on difference regression below,²²

$$\Delta s_{t+1} = \beta_1 + \beta_2(f_t - s_t) + \varepsilon_{t+1} \quad (3.1)$$

²² This is Eq. (2.7). It is reproduced in this chapter to ease references.

non-overlapping observation ensure that sampling frequency is equal to the maturity time of the forward contract (in this study it is 1-month forward contract/exchange rate), then ε_{t+1} will be serially uncorrelated.

The data sample spans the period January 1976 to December 1998 for the Euro-legacy currencies (French franc, German mark and Italian lira are involved in the consolidation of the European currencies into the Euro. The Euro currency begins on January 1999). For Canadian dollar and U.S. dollar, the monthly observations start from January 1976 to December 2005, while for Japanese yen the observation started from June 1978 to December 2005.

This study chooses G7 countries' currencies since they are heavily utilized in earlier studies of forward rate unbiasedness hypothesis (FRUH) (see Sarno, 2003; Engle, 1996). Furthermore, the sample period and the currencies used in this study are the same as Choi and Zivot (2007) and Sakoulis et al. (2010), except for Choi and Zivot (2007), the Japanese yen is not included.²³ As such, this study's findings are directly comparable to previous studies on a similar topic, in particular to Choi and Zivot (2007) and Sakoulis et al. (2010). Furthermore, using recent data does not change the econometric issues as recent findings still found evidences of the forward bias puzzle (see Bacheetta and Wincoop, 2010).

3.3 Forward Rate Unbiasedness Hypothesis (FRUH)

In Table 3.1 below, we report the ordinary least square (OLS) of difference regression (Eq. 2.7 or Eq. 3.1) as well as p -values for unbiasedness that implied both $\beta_1 = 0$ and $\beta_2 = 1$, joint test based on standard F -test. However, in most literatures, the primarily

²³ Refer to Table 2.3

focus is testing the restriction that $\beta_2 = 1$. The reason using this restriction is that it allows for the presence of a constant risk premium.²⁴ Based on difference regression, all currencies show that the joint hypothesis of $\beta_1 = 0$ and $\beta_2 = 1$ are rejected in all currencies at 5% and 1% significance level. The estimate of β_2 is also negative and statistically different than 1. Overall, the point estimates as well as the p -values of both tests confirm the usual findings of the forward discount being a biased predictor of the change in the future spot rate.

Table 3.1: Estimate of the difference regression

	Canadian dollar	French franc	German mark	Italian lira	Japanese yen	U.S. dollar
β_1	-0.004 (0.002)	-0.000 (0.002)	-0.005 (0.003)	0.005 (0.002)	-0.019 (0.005)	-0.004 (0.002)
β_2	-3.769 (0.922) [0.000]	-0.621 (0.647) [0.013]	-0.755 (0.703) [0.013]	-0.872 (0.518) [0.000]	-3.588 (1.077) [0.000]	-1.717 (0.763) [0.000]
R^2	0.051	0.008	0.005	0.019	0.028	0.018
Joint test: $H_0: \beta_1 = 0, \beta_2 = 1$	13.406 [0.000]	3.561 [0.030]	3.178 [0.043]	6.641 [0.002]	9.877 [0.000]	6.343 [0.002]

1) White heteroskedasticity-consistent standard errors in parenthesis.

2) p -values are in brackets for two tailed tests for $H_0: \beta_2 = 1$.

3) The joint test is based on F -test with corresponding p -values reported in brackets

In all cases, R^2 is very small, ranging from 0.005 in the case of German mark, to 0.051 for Canadian dollar. Thus, the results reported in Table 3.1 confirm the previous finding of forward bias puzzle based on our data sample of G7 currencies (see e.g. Bacchetta and Wincoop, 2010; Baillie and Bollerslev, 2000; Fama, 1984; Frankel and Poonawala, 2010). As noted by Froot (1990), based on 75 published articles on the puzzle, the average value of $\bar{\beta}_2$ estimated is -0.88 .

²⁴ The unbiasedness test is sometimes interpreted as a test for a time-varying risk premium (see Engel, 1996)

3.4 Descriptive Analysis of the G7 currencies

The forward discount is determined by subtracting the 1-month natural log of the forward exchange rate from the natural log of spot exchange rate. In Table 3.2, we report the summary statistics of forward discount of G7 currencies' used in this study.

Table 3.2: Descriptive statistics of forward discount

	Canadian dollar	French franc	German mark	Italian lira	Japanese yen	U.S. dollar
Mean	-0.0010	-0.0004	-0.0037	0.0029	-0.0046	-0.0019
Std. Dev	0.0019	0.0037	0.0026	0.0041	0.0016	0.0024
Skewness	0.4420	1.8037	-0.1221	1.2437	-0.7994	-0.5812
Kurtosis	7.0096	7.9189	3.655	5.065	3.9539	5.723
Jarque-Bera	252.88***	427.92***	5.62*	120.18***	47.80***	131.53***

1) ***, ** and * denotes significant level at 1%, 5% and 10% respectively.

Based on the descriptive statistics, three currencies' forward discount, Canadian dollar, French franc and Italian lira have right skewed distribution while the rest are having left skewed distribution. For Canadian dollar and French franc, they have long right tails compared to the others. Based on kurtosis, all forward discounts show Leptokurtic distribution where Canadian dollar and French franc's distributions are peaked relative to normal. The Jarque-Bera test shows that all forward discounts reject the normal distribution hypothesis.

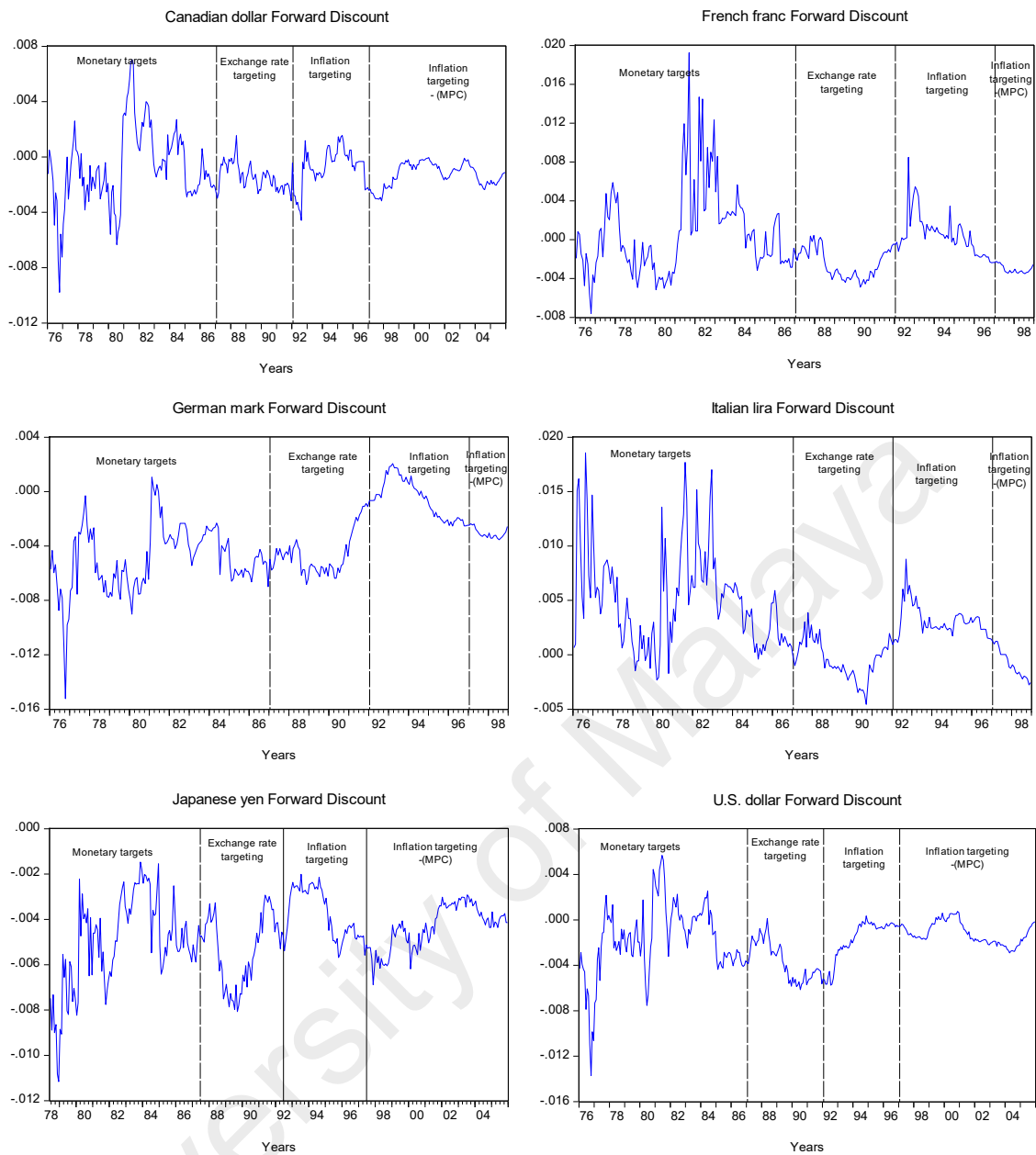


Figure 3.1: Time series plot of monthly forward discount of G7 currencies

In Figure 3.1 we plot the time series of forward discount, $(f_t - s_t)$, of G7 currencies that are used in this study. Since the exchange rate of G7 countries' currencies used in this study are quoted as foreign currency units (FCUs) per British pound (GBP), monetary policy of United Kingdom (UK) will influence most of the pattern in forward discount of all currencies used.

The UK's monetary policy can be divided into 4 phases that are relevant to our sample period. In Table 3.3, we report the phases and issues relevant to the monetary policy involved. Noticed that forward discount is much more volatile during monetary target policy of the UK.

Table 3.3: U.K.'s monetary policy from 1976 to presents

Monetary policy	Period (years)	Issues
Monetary targets	1976 – 1987	<ul style="list-style-type: none"> - Aimed to control various monetary aggregate in order to keep inflation down. - Significant global financial deregulation, including UK. - Relaxation of exchange and credit control in 1979 - Financial liberalization and innovation occur; as a result, monetary framework is unstable.
Exchange rate targets	1987 – 1992	<ul style="list-style-type: none"> - Policy makers in UK and Europe directed monetary policy to targeting the exchange rate - UK interest rate are set to keep the value of sterling within certain band relative to the German mark - Monetary authorities could not maintain the exchange rate target indefinitely - Some instability of both output and prices - Sharp depreciation of sterling in 1992
Inflation targeting- pre MPC (before the establishment of Monetary Policy Committee (MPC))	1992 – 1997	<ul style="list-style-type: none"> - UK exits from the Exchange Rate Mechanism - Sterling was allowed to float freely and a target for inflation was introduced - Targeting the rate of inflation as measured by the Retail Prices Index - Long term target of 2% inflation or less - Inflation fell this period - But inflation expectation remained above official target
Inflation targeting- MPC (after the establishment of Monetary Policy Committee (MPC))	1997 – presents	<ul style="list-style-type: none"> - Monetary Policy Committee (MPC) of the bank of England was given operational independence over monetary policy - MPC improved credibility and accountability leading to stable inflation

3.5 Summary

In this chapter, detail descriptions of the data used as sample are discussed. The sample and the currencies used are similar to Choi and Zivot (2007) and Sakoulis et al. (2010), except for Choi and Zivot (2007), where Japanese yen is not included. The finding of FRUH testing based on Fama's (1984) regression confirms the previous finding of negative coefficient. Since the currency used are quoted as foreign currency unit (FCUs) per British pound (GBP), the monetary policy of the UK might influence most of the pattern in forward discount of all currencies used.

University of Malaya

Chapter 4: Long Memory of Forward Discount

4.1 Introduction

In this chapter, the discussion aims to answer the first objective of this study, namely long memory.²⁵ Based on this objective, the long memory parameter will be estimated using the more efficient semiparametric method before and after adjusting for structural breaks in the forward discount. Lastly, two formal statistical tests of spurious long memory will be conducted to complement the informal technique of structural break adjustment suggested by Choi and Zivot (2007).

4.2 Past Studies on Long Memory

Studies conducted by Baillie and Bollerslev (1994b, 2000), Sakoulis and Zivot (2001) and Maynard and Phillips (2001) have drawn considerable attention with their findings regarding long memory in the forward discount. They concluded that the forward discount is fractionally integrated, hence rejecting the two extreme cases of the stationary $I(0)$ and the non-stationary $I(1)$ levels of the series. The series has a long memory property represented by its hyperbolically decayed autocorrelation, or the periodogram shows a peak around zero frequency. This important finding provides the idea of unbalanced regression, as Baillie and Bollerslev (1994b, 2000) noted, which invalidates the standard statistical inference of regression-based testing of the forward rate unbiasedness hypothesis as discussed earlier. Thus, they argued that the forward bias puzzle exists mainly due to the statistical properties of the data.

²⁵ Refer to section 1.4 for detail discussion.

Despite its appealing findings regarding long memory in the forward discount as the possible cause of the puzzle, however, little attention has been given to the possibility of spurious long memory in the forward discount. Structural breaks or regime switching in a series may contribute to long memory, as discussed in Granger and Hyung (2004), Diebold and Inoue (2001), Granger and Hyung (1999) and Granger (1999). The importance of distinguishing between the true or spurious long memory is due to fact that statistical inference is quite different between these two processes (see Shao, 2011). Given that structural breaks or regime switching might occur due to various reasons (e.g. the intervention of monetary authorities in business cycles), therefore, the validity of earlier findings on long memory in the forward discount is questionable.

One study that addresses the issue of spurious long memory in forward discount was conducted by Choi and Zivot (2007). In the study, Choi and Zivot (2007) estimate the long memory parameter using semiparametric approaches twice; before and after adjusting for structural breaks. In estimating the long memory parameter, they used modified log periodogram (MLP) of Phillips (2007) and Kim and Phillips (2000, 2006), while for structural breaks in mean they rely on Bai and Perron's (2003) model. By using monthly data of G7 countries' currencies in terms of U.S. dollars that cover the period from 1976:1 to 1991:1, they provide evidence of less persistent process in forward discount after considering break in mean of the series.²⁶ However, they still found long memory in forward discount even after removing structural breaks. Even though this finding is appealing, study conducted by Choi and Zivot (2007) does not provide any statistical prove that the process is indeed true long memory process. It was assumed that the long memory process in forward discount is spurious due to structural breaks. Furthermore, discriminating between structural breaks and long memory is

²⁶ However, Japanese yen is excluded from their study due to different sample period.

difficult because structural breaks tests are often biased toward rejection of the null hypothesis of no breaks when the process is indeed long memory (see Qu, 2011).

Distinguishing between the true long memory processes against the spurious process one is a difficult task. Many studies show that the time series with structural breaks can induce a strong persistence in the autocorrelation function (ACF) (it exhibits a slow rate of decay) and hence generate spurious long memory process. The difficulties involved in the determining whether it is a true long memory process or due to structural breaks is shared in studies conducted by Perron and Qu (2010), Granger and Hyung (2004), Diebold and Inoue (2001) and Granger (1999). Recently, Ohanissian et al. (2008) developed a test to determine whether the series has a true or spurious long memory. However, their test required Gaussianity and the bandwidth of the Geweke and Porter-Hudak (1983) (GPH) estimate of long memory in the test was stringent and required a large sample size to be useful.²⁷ Furthermore, the test was based on the idea of temporal aggregation which is not compelling given that the forward discount is simply an interest rate differential.²⁸

Other recently developed test of spurious long memory is done by Qu (2011) and Shimotsu (2006). Qu (2011) proposes a tests statistic for the null hypothesis that a given time series is a stationary long memory process against the alternative hypothesis that it is affected by regime change or a smoothly varying trend. This test is based on profiles likelihood function of the local Whittle estimator (Robinson, 1995). Although the estimation is based on semiparametric technique where it does not require determining the short memory process, he offers 'prewhitening' to control the test size in the presence of short memory. The test also includes a trimming parameter that ensures a reliable asymptotic approximation even in a small sample.

²⁷ The test is conducted to analyze 5-minute returns with around 600,000 observations.

²⁸ Refer to Eq. (2.4).

Shimotsu (2006) offers two simple tests based on certain time domain properties of $I(d)$ process. In the first test, if the time series process is not affected by regime change or true $I(d)$ process, then each subsample of the time series must also follow an $I(d)$ process with the same value of d . The test is conducted using adjusted Wald statistic by comparing the estimated of d for the entire time period with the estimate of d for subperiods. If null of constancy in parameter d is rejected, the time series might be affected by regime change. In the second test, if a time series process follows an $I(d)$ process, then its d th-differenced series will follow an $I(0)$ process. The standard Kwiatkowski et al. (1992) (KPSS) test and Phillips-Perron tests on the d -differenced series can determine whether the process is true long memory process due to the fact that a spurious long memory process will fail to be $I(0)$ process even after d -differenced.

The focus of this chapter is to extend the study by Choi and Zivot (2007). Firstly, we replicate the method used in Choi and Zivot (2007) where we add the estimation of long memory parameter before and after adjusting for structural breaks based on two-step feasible exact local Whittle (FELW) of Shimotsu and Phillips (2006). There are two advantages of local Whittle over GPH estimator. One, the local Whittle estimator is more efficient than GPH estimator is (see Robinson, 1995). Two, it is robust with respect to heteroskedasticity of a certain degree (Robinson and Henry, 1999; Shao and Wu, 2007). It should be noted that modified log periodogram (MPL) that is used in Choi and Zivot (2007) is just the modification of GPH estimator. Furthermore, FELW estimation also has the ability to accommodate both stationary and nonstationary $I(d)$ process in a single data generating process (DGP) (see Shimotsu and Phillips, 2006). Finally, we report the result of formal statistical testing of spurious long memory by Shimotsu (2006) and Qu (2011).

4.3 Methodologies

4.3.1 ARFIMA and Classification of Long Memory Parameter d

A covariance stationary process, X_t , has a long memory process if:

$$\sum_{k=-n}^n |\rho(k)| \rightarrow \infty, \quad \text{as } n \rightarrow \infty \quad (4.1)$$

where $\rho(\cdot)$ representing the autocorrelation coefficient of the stationary long memory process satisfies

$$\rho(k) \propto k^{2d-1} \quad k \rightarrow \infty \quad 0 < d < \frac{1}{2} \quad (4.2)$$

where it is hyperbolically decayed. This definition fits a model called the fractionally integrated autoregressive moving average (ARFIMA), a class of autoregressive moving average (ARMA), which allows d assuming any real value. This model is considered an intermediate situation between a short memory ARMA and a fully integrated ARIMA, which is restricted to the integer domain only. This framework is attributed to Hosking (1981) and Granger and Joyeux (1980).

A time series $\mathbf{X} = \{X_1, \dots, X_T\}$ follows an ARFIMA(p, d, q) process if

$$\Phi(L)(1-L)^d(X_t - \mu) = \Theta(L)\varepsilon_t, \quad \varepsilon_t \sim iid(0, \sigma^2) \quad (4.3)$$

where L is the backward-shift operator, $\Phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$, $\Theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$ is an autoregressive and moving average polynomial, respectively, with unit roots that are outside the unit circle. $(1-L)^d$ is the fractional differencing operator defined as

$$(1-L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)} \quad (4.4)$$

a binomial expansion where $\Gamma(\cdot)$ is a gamma function. The value of parameter d can fall into five categories, where:²⁹

- a) $d \in (-0.5, 0)$: anti-persistent memory
- b) $d \in (0, 0.5)$: stationary long memory
- c) $d \in [0.5, 1)$: nonstationary long memory but with mean reverting and finite impulse response weights³⁰
- d) $d \geq 1$: nonstationary
- e) $d = 0$: short memory³¹

4.3.2 Semiparametric Estimation of Parameter \hat{d}

There are several approaches to estimating the parameter above. In semiparametric methods, two popular methods are log-periodogram or GPH estimator proposed by Geweke and Porter-Hudak (1983) and local Whittle estimator of Robinson (1995). Let us consider covariance stationary long memory process where we assume the spectral density $f(\lambda)$ of the process X_t satisfies

$$f(\lambda) \sim G\lambda^{-2d}, \quad \text{as } \lambda \rightarrow 0 + \quad (4.5)$$

The most widely used long memory process is a fractionally integrated process X_t , in form of the model:

$$(1 - L)^d X_t = \varepsilon_t \quad (4.6)$$

where ε_t is a covariate stationary process. Define the discrete Fourier transform (DFT)

as

²⁹ For the classification of d in ARFIMA, see Baillie (1996) and Baillie and Bollerslev (1994b, 2000).

³⁰ Parameter d in between 0.5 and 1 is being mistakenly viewed as a unit root non-stationary process. The distinction between unit root non-stationary process and nonstationary long memory with mean reverting is significant where the persistence shocks in both situations are radically different (see Kapetanios and Shin, 2011).

³¹ Stationary and invertible ARMA.

$$w_X(\lambda_j) = (2\pi T)^{-1/2} \sum_{t=1}^T X_t \exp(i\lambda_j t), \quad \lambda_j = \frac{2\pi j}{T}, j = 1, \dots, m < T \quad (4.7)$$

and the periodogram ordinates of

$$I_X(\lambda_j) = |w_X(\lambda_j)|^2 \quad (4.8)$$

The idea of GPH is to estimate the spectral density by the periodogram and to take logarithm on both sides of the above equation. This gives a linear regression model in the memory parameter, which can be estimated by least squares (LS). The estimator is given by 1/2 times the least squares estimator of the slope parameter in the regression of $\{\log I_X\}$ on a constant and the regressor variable

$$Y_j = \log|1 - \exp(-i\lambda_j)| = \frac{1}{2} \log(2 - 2 \cos \lambda_j) \quad (4.9)$$

By definition, the GPH estimator based on the first m periodogram ordinates

$$\hat{d}_{GPH} = \frac{0.5 \sum_{j=1}^m (Y_j - \bar{Y}) \log I_X(\lambda_j)}{\sum_{j=1}^m (Y_j - \bar{Y})^2} \quad (4.10)$$

where $\bar{Y} = \frac{1}{m} \sum_{j=1}^m Y_j$. One of the main issues with the GPH estimator is the selection of periodogram ordinates m when ε is autocorrelated. As Agiakloglou et al. (1993) showed, the GPH estimator is biased and inefficient when ε_t is the $AR(1)$ or $MA(1)$ process, where either has a large parameter. In the context of forward discount, Sakoulis et al. (2010) showed that the forward discount is indeed a highly persistent $AR(1)$ process. Furthermore, the GPH estimator is inconsistent for $d > 1$.

Velasco (1999) showed that consistency only applies for $\frac{1}{2} < d < 1$ with Gaussian assumptions. According to Kim and Phillips (2006), in the presence of the unit root, the estimate is not consistent when $d > 1$ and is consistent for $\frac{1}{2} < d < 1$ without Gaussian assumptions, where $d = 1$ is the boundary limit.

Phillips (2007) and Kim and Phillips (2000, 2006) addressed the issue of consistency in the GPH estimate by suggesting what is known as the modified log periodogram (MLP) regression estimator, which looks at $d = 1$ or the unit root. The estimator they suggested is consistent for both $d \leq 1$ and $d > 1$. It applies a similar regression, but with different periodogram ordinates. The suggested periodogram ordinates are $I_v(\lambda_j) = v_X(\lambda_j)v_X(\lambda_j)^*$ where the modification is expressed as

$$v_X(\lambda_j) = w_X(\lambda_j) + \frac{\exp(i\lambda_j)}{1 - \exp(i\lambda_j)} \frac{X_T}{\sqrt{2\pi T}} \quad (4.11)$$

where $v_X(\lambda_j)$ is the modified discrete Fourier transform (DFT), which results in. Thus, the modified long parameter estimate becomes

$$\tilde{d}_{MLP} = \frac{0.5 \sum_{j=1}^m (Y_j - \bar{Y}) \log I_v(\lambda_j)}{\sum_{j=1}^m (Y_j - \bar{Y})^2} \quad (4.12)$$

As shown in Kim and Phillips (2000, 2006), the distribution of \tilde{d}_{MLP} follows

$$\sqrt{m}(\tilde{d}_{MLP} - d) \xrightarrow{d} \mathcal{N}\left(0, \frac{\pi^2}{24}\right) \quad (4.13)$$

which allow us to compute the confidence intervals.

For local Whittle estimation, we start with the following Gaussian objective function defined in terms of the parameter d and G

$$Q_m(G, d) = \frac{1}{m} \sum_{j=1}^m \left[\log(G\lambda_j^{-2d}) + \frac{\lambda_j^{2d}}{G} I_X(\lambda_j) \right], \quad j = 1, \dots, m < T \quad (4.14)$$

The local Whittle procedure estimates G and d by minimizing $Q_m(G, d)$ such that

$$(\hat{G}, \hat{d}) = \arg \min_{G \in (0, \infty), d \in [\Delta_1, \Delta_2]} Q_m(G, d) \quad (4.15)$$

where Δ_1 and Δ_2 are numbers such that $-1/2 < \Delta_1 < \Delta_2 < \infty$. By concentrating (4.14) with respect to G gives

$$\hat{d}_W = \arg \min_{d \in [\Delta_1, \Delta_2]} R(d) \quad (4.16)$$

where

$$R(d) = \log \hat{G}(d) - 2d \frac{1}{m} \sum_1^m \log \lambda_j, \quad \hat{G}(d) = \frac{1}{m} \sum_1^m \lambda_j^{2d} I_X(\lambda_j) \quad (4.17)$$

Shimotsu and Phillips (2006) suggest the two-step feasible exact local Whittle (FELW)

where the objective function of

$$R_F(d) = \log \hat{G}_F(d) - 2d \frac{1}{m} \sum_1^m \log \lambda_j, \quad \hat{G}_F(d) = \frac{1}{m} \sum_1^m I_{\Delta^d(x - \tilde{\mu}(d))}(\lambda_j) \quad (4.18)$$

where $I_{\Delta^d(x - \tilde{\mu}(d))}(\lambda_j)$ denotes the periodogram of $(1 - L)^d(X_t - \tilde{\mu}(d))$ and defined by

$$\tilde{\mu}(d) = w(d)\bar{X} + (1 - w(d))\bar{X}_1 \quad (4.19)$$

where $w(d)$ is a twice continuously differentiable weight function such that $w(d) = 1$ for $d \leq 1/2$ and $w(d) = 0$ for $d \geq 3/4$. The FELW estimator is defined as

$$\hat{d}_{FELW} = \tilde{d} - R_F(\tilde{d})' / R_F(\tilde{d})'' \quad (4.20)$$

where \tilde{d} is a first-stage $m^{1/2}$ consistent estimator of d . Following Hassler and Meller (2014), the distribution is

$$2\sqrt{m}(\hat{d}_{FELW} - d) \xrightarrow{d} \mathcal{N}(0,1) \quad (4.21)$$

which allow us to compute the approximate confidence intervals.

4.3.3 Bai and Perron's (2003) Structural Breaks Model

This section is based on Bai and Perron's (2003), which is used to determine the multiple breaks in the mean of the forward discounts. The model is defined as follows:

$$y_t = z_t' \delta_j + u_t, \quad t = T_{j-1} + 1, \dots, T_j \quad (4.22)$$

for $j = 1, \dots, m + 1$, $T_0 = 0, T_{m+1} = T$, where T_j indices represent the break points, which are explicitly treated as unknown. In this chapter, y_t represents the forward discount ($f_t - s_t$), while δ_j is the mean of the forward discount in different regimes. This can be achieved by applying a constant as the regressor ($z_t = 1$). The estimate of δ_j is obtained by minimizing the sum of squared residuals, as follows:

$$(Y - \bar{Z}\delta)'(Y - \bar{Z}\delta) = \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} (y_t - z_t' \delta_i)^2 \quad (4.23)$$

where $\bar{Z} = \text{diag}(Z_1, \dots, Z_{m+1})$. Let $\hat{\delta}(\{T_1, \dots, T_m\})$ denote an estimate based on the m -partition (T_1, \dots, T_m) , and substituting this in the above denotes the resulting sum of squared residuals as $S_T(T_1, \dots, T_m)$; the estimate break points are defined as follows:

$$(\hat{T}_1, \dots, \hat{T}_m) = \underset{T_1, \dots, T_m}{\text{argmin}} S_T(T_1, \dots, T_m) \quad (4.24)$$

where minimization occurs over all the m -partition (T_1, \dots, T_m) . This results in breakpoint estimators being global minimizers of the objective function. To compute global minimizers, Bai and Perron (2003) suggest using dynamic programming, which is not computationally excessive compared to grid searches.

Two methods of selecting number of breaks reported are Bayesian information criterion (BIC) and sequential procedure. As Bai and Perron (2003) argued, the sequential procedure is superior in the presence of serial correlation in the errors, where BIC overestimate the true number of breaks.

However, the sequential procedure may break down in the presence of multiple breaks, where it is difficult to reject the null hypothesis break of zero versus one break. This may occur when two changes are present and the coefficient returns to its original

value after the second break.³² However, a different scenario exists when it comes to the null hypothesis of zero breaks versus higher breaks. In the event the sequential procedure breakdown, the first step is to determine the existence of breaks based on the double maximum test (*UDmax* and *WDmax*). Once the test result suggests breaks do occur, then the number of breaks can be decided by using the combination of BIC, $supF_T(\ell)$ and the $supF_T(\ell + 1|\ell)$ where we select m such that the $supF_T(\ell)$ and $supF_T(\ell + 1|\ell)$ is insignificant for $\ell \geq m$.

4.3.4 Choi and Zivot's (2007) Structural Break Adjustments

Choi and Zivot (2007) introduced a technique to incorporate breaks in the series to mimic the effects of spurious long memory. To adjust for structural breaks, the estimate of the parameter \hat{d}_{MLP} and \hat{d}_{FELW} using the residual series $\hat{u}_t = (f_t - s_t) - \hat{\delta}_j$ where $\hat{\delta}_j$ represents the mean of the forward discount from Bai and Perron's (2003) tests for $j = 1, \dots, m$ breaks. This technique allows us to analyse the existence of the true long memory component after taking into account the breaks in the mean of the forward discount.

4.3.5 Shimotsu's (2006) 'Simple but Effective' Tests

The first test suggested by Shimotsu (2006) is to compare the estimate of \hat{d} for the entire time period with the estimate of \hat{d} for sub periods based on the adjusted Wald statistics. More formally, the null hypothesis is $H_0: d_0 = d_{0,1} = \dots = d_{0,b}$, where d_0 is the true value of d for the entire sample, and $d_{0,i}$ is the true value of d from the i^{th}

³² See Bai and Perron (2003) for detail discussion

subsample. Using the two-step feasible exact local Whittle (FELW) estimator \hat{d} of d_0 , the adjusted Wald statistic for testing H_0 is

$$W_c = 4m \left(\frac{c_{m/b}}{m/b} \right) A \hat{d}_b (A \Omega A')^+ (A \hat{d}_b)' \quad (4.25)$$

where

$$\hat{d}_b = \begin{pmatrix} \hat{d} - d_0 \\ \hat{d}^{(1)} - d_0 \\ \vdots \\ \hat{d}^{(b)} - d_0 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & -1 \end{pmatrix} \quad (4.26)$$

$(A \Omega A')^+$ denotes a generalised inverse of $A \Omega A' = bI_b - i_b i_b'$ (where I_b is the identity matrix of order b and i_b is the unit b -vector), and c_m is a small-sample correction factor to allow for the large size of the finite sample variance:

$$c_m = \sum_{i=1}^m v_i^2, \quad v_i = \log i - \frac{1}{m} \sum_{i=1}^m \log i \quad (4.27)$$

Under H_0 , W_c has an asymptotic χ^2 distribution with $T - 1$ degrees of freedom and the usual decision criterion applies.

The second test is based on the behaviour of the d -differenced series. If an $I(d)$ series is differenced \hat{d} times, where \hat{d} is a consistent estimate of d through FELW, the resultant should be $I(0)$. Shimotsu considers two tests: Phillips and Perron's (1998) Z_t for the partial sum process of the differenced series, and the KPSS test for the differenced series. To allow for bias caused by the short-run dynamics of the series, the differencing is carried out on a mean-adjusted series as defined in Eq. (4.19).

4.3.6 Qu's (2011) Spurious Long Memory Test

Qu (2011) proposes a test of spurious long memory with the null hypothesis that a given time series is a stationary long memory process against the alternative hypothesis that it is affected by regime change or a smoothly varying trend. The test statistics is:

$$W = \sup_{r \in [\varepsilon, 1]} \left(\sum_{j=1}^m v_j^2 \right)^{-1/2} \left| \sum_{j=1}^{[mr]} v_j \left(\frac{\lambda_j^{2\hat{d}_w}}{\hat{G}(d)} I_X(\lambda_j) - 1 \right) \right| \quad (4.28)$$

based on local Whittle likelihood function of Robinson (1985) as defined in Eq. (4.15) and (4.16). The term v_j is defined as $v_j = \log \lambda_j - (1/m) \sum_{j=1}^m \log \lambda_j$. The test offers ‘prewhitening’ to control the present of short memory. The trimming parameter of ε ensure a reliable asymptotic approximation even in small sample. Qu (2011) suggests choosing $\varepsilon = 0.05$ for a small sample.

4.4 Autocorrelation Function and Periodogram

An informal approach to determine whether a series has a long memory property is to plot their autocorrelation function (ACF). Figure 4.1 shows each G7 countries’ autocorrelation of the forward discount ($f_t - s_t$), return ($s_{t+1} - s_t$) and spot exchange rate (s_t). Based on the definition of long memory, the autocorrelations of long memory process should lie between the autocorrelations of a stationary autoregressive process and non-stationary process. As shown in Figure 4.1, all G7 currencies’ forward discount lies in between the spot exchange rate and return. Similarly, the periodogram shows a peak around zero frequency in Figure 4.2. However, as mention earlier structural breaks can induce a strong persistence in the autocorrelation function (ACF)(it exhibits a slow rate of decay) and hence generate spurious long memory process

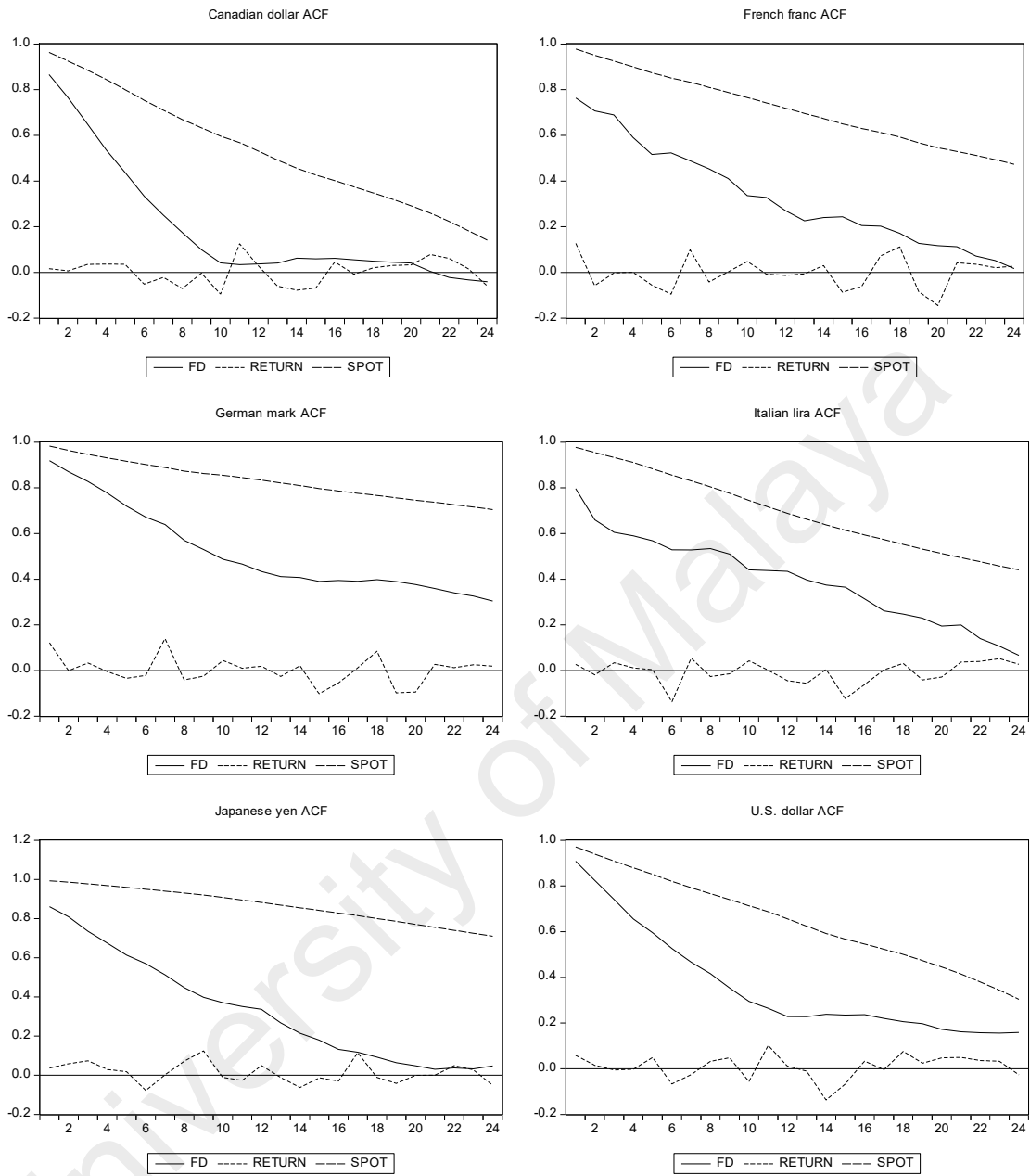


Figure 4.1: Autocorrelation function (ACF) of G7 currencies

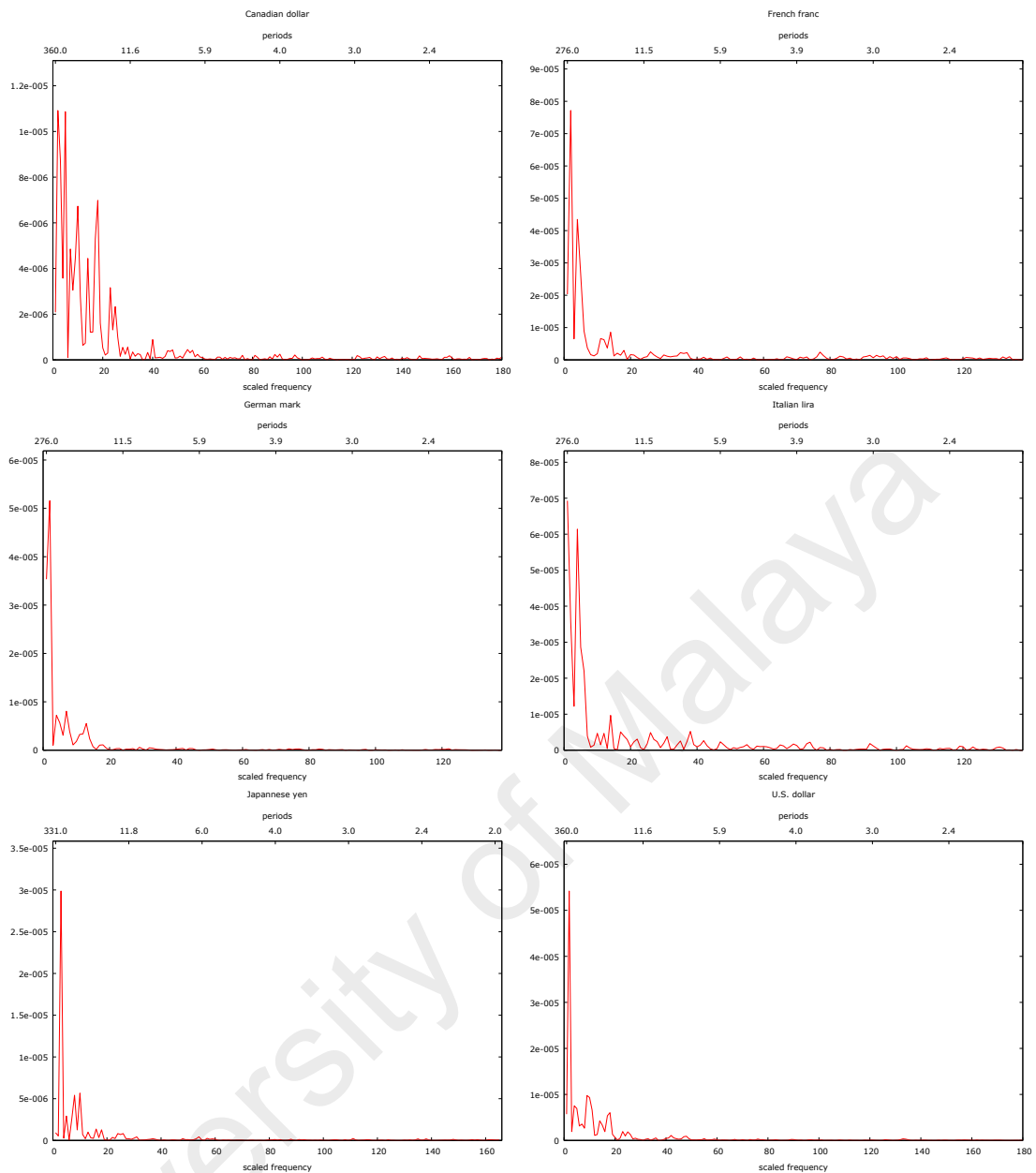


Figure 4.2: Periodogram of forward discount³³

4.5 Findings

4.5.1 Semiparametric Estimates before Break Adjustments³⁴

In Table 4.1 and 4.2, we report the two semiparametric estimations of MLP and FELW before adjustments for structural breaks. The periodogram ordinates m for MLP is selected following Choi and Zivot (2007), while for FELW as suggested by Hassler and Meller (2014). Most of the estimates are greater than 0.5 except for Italian lira at $m =$

³³ The periodogram is plot-using Gretl, an open-source econometric package.

³⁴ MLP estimate is conducted using STATA module of MODLPR written by Christopher F. Baum. For FELW estimate, MATLAB codes written by Katsumi Shimotsu is used and available at http://shimotsu.web.fc2.com/Site/Matlab_Codes.html

$\{T^{0.70}, T^{0.75}\}$ and Canadian dollar at $m = \{T^{0.60}\}$ based on MLP and FELW estimator respectively. As such, most of G7's forward discount is nonstationary long memory process with the exception to Italian lira and Canadian dollar, where evidences of stationary long memory process are found at some level of m .

The results for Canadian dollar are quite similar to those of Choi and Zivot (2007), They reported that \hat{d}_{MLP} at $m = \{T^{0.70}\}$ is 0.619, while ours is 0.648. However, when comparing with FELW estimation, \hat{d}_{FELW} for Canadian dollar is 0.773 for the same m . In contrast, Maynard and Phillips (2001) reported a much higher result for the Canadian dollar, at $m = \{T^{0.75}\}$, it is 0.957 based on MLP estimation. This might be due to different sample period as in Maynard and Phillips (2001), the sample from 1986:11 to 1998:3.

In addition, study by Baillie and Bollerslev (1994b) found $\hat{d} = 0.45$ for the same currency. It should be noted that in Baillie and Bollerslev (1994b), the long memory parameter of *ARFIMA* model is estimated through parametric method of exact maximum likelihood estimation (*EML*) of Sowell (1992).

In particular, using parametric estimation, the researcher needs to specify the short memory or *ARMA* process, which is the order of the autoregressive (p) and the moving average (q) polynomials of the *ARFIMA*(p, d, q) model. As in Baillie and Bollerslev (1994b), they model the forward discount as *ARFIMA*(2, d , 0) model. However, Sowell's (1992) *EML* estimation requires a stationary univariate *ARFIMA* model, and if the series is nonstationary, it requires differencing before estimation.

In Baillie and Bollerslev (1994b), the forward discount is difference before applying the Sowell's (1992) *EML* estimation, with the assumption that forward discount is indeed nonstationary. Given that nonstationarity of forward discount is uncertain as discussed in Chapter 2, one can avoid this by relying on semiparametric. Furthermore, with semiparametric, one does not need to specify the *ARMA* process.

French franc estimation of long memory based on *MLP* is also quite similar to Choi and Zivot (2007). It was reported in Choi and Zivot (2007) that \hat{d}_{MLP} is 0.615 and ours is 0.691. Even though we used the same sample period as Choi and Zivot (2007), in their study the currencies used are quoted as foreign currency units (FCUs) per U.S. dollar, while ours is per British pound (GBP). This might explain some variation in estimation results. However, based on *FELW*, French franc reported slightly higher persistence of $\hat{d}_{FELW}=0.747$ at $m = \{T^{0.70}\}$. Yet again, Maynard and Phillips (2001) reported a much higher result for French franc, at $m = \{T^{0.75}\}$, it is 0.885.

The result for German mark reported slightly more persistence as compare to Choi and Zivot (2007) and Baillie and Bollerslev (1994b). In Choi and Zivot (2007) and Baillie and Bollerslev (1994b), the estimate of \hat{d} is 0.725 and 0.77 respectively. However, it is slightly less persistent as compared to Maynard and Phillips (2001), where it was reported as 0.888 while ours is 0.809 for German mark. Interestingly, based on *FELW*, the result is not much different to Maynard and Phillips (2001). For German mark, \hat{d}_{FELW} is 0.875.

The Italian lira shows somewhat less persistence process as compared to Choi and Zivot (2007). In Choi and Zivot (2007), \hat{d}_{MLP} is 0.536 and ours is 0.411. Though, it is quite similar to *FELW* result of 0.592. In general, most of the currencies used show evidence of nonstationary long memory based on both estimation methods, except for Italian lira. Interestingly, Italian lira's 90% confidence interval includes a stationary long memory range of $\hat{d} < 0.5$ based on *MLP* and *FELW*. For German mark, the 90% confidence interval includes the nonstationary and non-mean reverting range of $\hat{d} \geq 1$.

4.5.2 Results of The Structural Breaks Model³⁵

Based on the results from Table 4.3, for the Canadian dollar, the *UDmax* and *WDmax* tests suggest no break for the whole sample period. All $SupF_T(1)$ through $SupF_T(5)$ are insignificant at 5% level for this currency. Similar finding by $SupF_T(l + 1|l)$ test where all are insignificant for any level of l with the sequential test also suggests zero break. Even though the BIC suggests two breaks, the rest of the tests are in favour of zero breaks. Thus, for Canadian dollar, no break is selected.

In the case of French franc, the sequential test shows no break in the sample. However, the *UDmax* and *WDmax* tests suggests otherwise. This may signal the breakdown of sequential test as mentioned earlier. The BIC and $SupF_T(l + 1|l)$ tests are in favor of 4 breaks for French franc. Thus, for French franc, 4 breaks are selected.

For German mark, Italian lira, Japanese yen and U.S. dollar, the *UDmax* and *WDmax* tests suggests the existence of multiple breaks in the sample. This further supported by BIC where 5 breaks are selected based on these tests. The results of $SupF_T(5)$ of these currencies are all significant for these currencies. Thus, for German mark, Italian lira, Japanese yen and U.S. dollar, 5 breaks are selected for the sample period.

The number of breaks is consistent for French franc and the German mark as reported in Choi and Zivot (2007) where they also found 4 and 5 breaks respectively. For Italian lira, we found 5 breaks while in Choi and Zivot (2007) it is only 4. However, interestingly, we fail to find any break for Canadian dollar where in Choi and Zivot (2007), 3 breaks are found for the same currency. Nevertheless, the variances in number of breaks in this study as compared to Choi and Zivot (2007) might be due different foreign currency units (FCUs) quoted as mentioned in the previous section.

³⁵ Structural breaks test is conducted using GAUSS code written by Pierre Perron, available at <http://people.bu.edu/perron/>

Table 4.1: Estimate of \hat{d}_{MLP} without adjustment for breaks

	m	\hat{d}_{MLP}	S.E.	90% C.I.
Canadian dollar	$T^{0.70}$	0.648	0.096	[0.490, 0.806]
	$T^{0.75}$	0.762	0.081	[0.629, 0.895]
	$T^{0.80}$	0.804	0.071	[0.687, 0.920]
French franc	$T^{0.70}$	0.691	0.109	[0.511, 0.871]
	$T^{0.75}$	0.757	0.091	[0.607, 0.907]
	$T^{0.80}$	0.655	0.080	[0.523, 0.787]
German mark	$T^{0.70}$	0.809	0.091	[0.660, 0.958]
	$T^{0.75}$	0.892	0.077	[0.766, 1.018]
	$T^{0.80}$	0.789	0.067	[0.680, 0.899]
Italian lira	$T^{0.70}$	0.411	0.107	[0.235, 0.587]
	$T^{0.75}$	0.361	0.082	[0.225, 0.497]
	$T^{0.80}$	0.502	0.085	[0.362, 0.641]
Japanese yen	$T^{0.70}$	0.608	0.091	[0.458, 0.758]
	$T^{0.75}$	0.760	0.085	[0.620, 0.901]
	$T^{0.80}$	0.725	0.071	[0.609, 0.842]
U.S. dollar	$T^{0.70}$	0.691	0.092	[0.541, 0.842]
	$T^{0.75}$	0.777	0.080	[0.645, 0.908]
	$T^{0.80}$	0.857	0.074	[0.736, 0.978]

1) Selection of m following Choi and Zivot (2007)

2) \hat{d}_{MLP} is from Eq. (4.12)

3) 90% confidence intervals (CI) are constructed from Eq. (4.13)

Table 4.2: Estimate of \hat{d}_{FELW} without adjustment for breaks

	m	\hat{d}_{FELW}	S.E	90% C.I.
Canadian dollar	$T^{0.60}$	0.497	0.086	[0.356, 0.638]
	$T^{0.65}$	0.725	0.074	[0.604, 0.846]
	$T^{0.70}$	0.773	0.064	[0.668, 0.878]
French franc	$T^{0.60}$	0.753	0.093	[0.601, 0.905]
	$T^{0.65}$	0.647	0.080	[0.515, 0.779]
	$T^{0.70}$	0.747	0.070	[0.632, 0.862]
German mark	$T^{0.60}$	0.798	0.093	[0.646, 0.950]
	$T^{0.65}$	0.897	0.080	[0.765, 1.029]
	$T^{0.70}$	0.875	0.070	[0.760, 0.990]
Italian lira	$T^{0.60}$	0.659	0.093	[0.507, 0.811]
	$T^{0.65}$	0.602	0.080	[0.470, 0.734]
	$T^{0.70}$	0.592	0.070	[0.477, 0.707]
Japanese yen	$T^{0.60}$	0.692	0.088	[0.548, 0.836]
	$T^{0.65}$	0.810	0.076	[0.685, 0.935]
	$T^{0.70}$	0.752	0.066	[0.644, 0.860]
U.S. dollar	$T^{0.60}$	0.660	0.086	[0.519, 0.801]
	$T^{0.65}$	0.720	0.074	[0.599, 0.841]
	$T^{0.70}$	0.771	0.064	[0.666, 0.876]

1) Selection of m as suggested by Hassler and Meller (2014)

2) \hat{d}_{FELW} is from Eq. (4.20)

3) 90% confidence intervals (CI) are constructed from Eq. (4.21)

Turning to Table 4.4, all forward discounts have two breaks during the period of monetary targeting of 1976 to 1987. As discussed in Table 3.3, this is the most volatile period where significant global financial deregulation takes place. Couple with financial liberalization and innovation, monetary framework during this period is unstable. For European countries' currencies in this study, French franc, German mark and Italian lira, they experience another two breaks during the period of exchange rate targeting from 1987 to 1992. These two breaks occur because during this period, European monetary policy are targeting the exchange rate and maintaining this policy result in some instability in both output and prices.

Table 4.3: Structural break test for forward discount

	Canadian dollar	French franc	German mark	Italian lira	Japanese yen	U.S. dollar
Sequential	0	0	0	1	1	0
BIC	2	4	5	5	5	5
<i>UDmax</i>	3.84	21.50*	28.54*	14.57*	13.63*	9.85*
<i>WDmax</i>	6.07	35.46*	49.08*	23.82*	27.56*	11.70*
<i>SupF_T</i> (1)	3.84	5.284	4.81	14.57*	10.09*	1.89
<i>SupF_T</i> (2)	2.82	21.502*	24.99*	12.05*	13.63*	9.85*
<i>SupF_T</i> (3)	2.51	15.60*	17.75*	10.95*	6.40*	4.62
<i>SupF_T</i> (4)	2.66	19.08*	28.54*	12.68*	6.13*	3.90
<i>SupF_T</i> (5)	2.77	16.16*	21.60*	10.85*	12.56*	4.32*
<i>SupF_T</i> (2 1)	2.96	19.39*	17.04*	3.79	5.49	2.45
<i>SupF_T</i> (3 2)	2.09	3.02	4.13	0.86	2.50	1.86
<i>SupF_T</i> (4 3)	1.79	26.94*	12.04*	11.05	4.84	0.22
<i>SupF_T</i> (5 4)	-	0.16	-	0.01	37.60*	0.28

1) * denotes 5% level of significant.

During inflation targeting pre and post MPC period, most of the currencies experience one break respectively. This might be contributed by the introduction of independence institution of Monetary Policy Committee (MPC), which improved credibility and accountability that lead toward stable inflation. The time series plots of forward discount with multiple breaks in mean line are shown in Figure 4.3 below. The mean lines are taken from the result as reported in the Table 4.4.

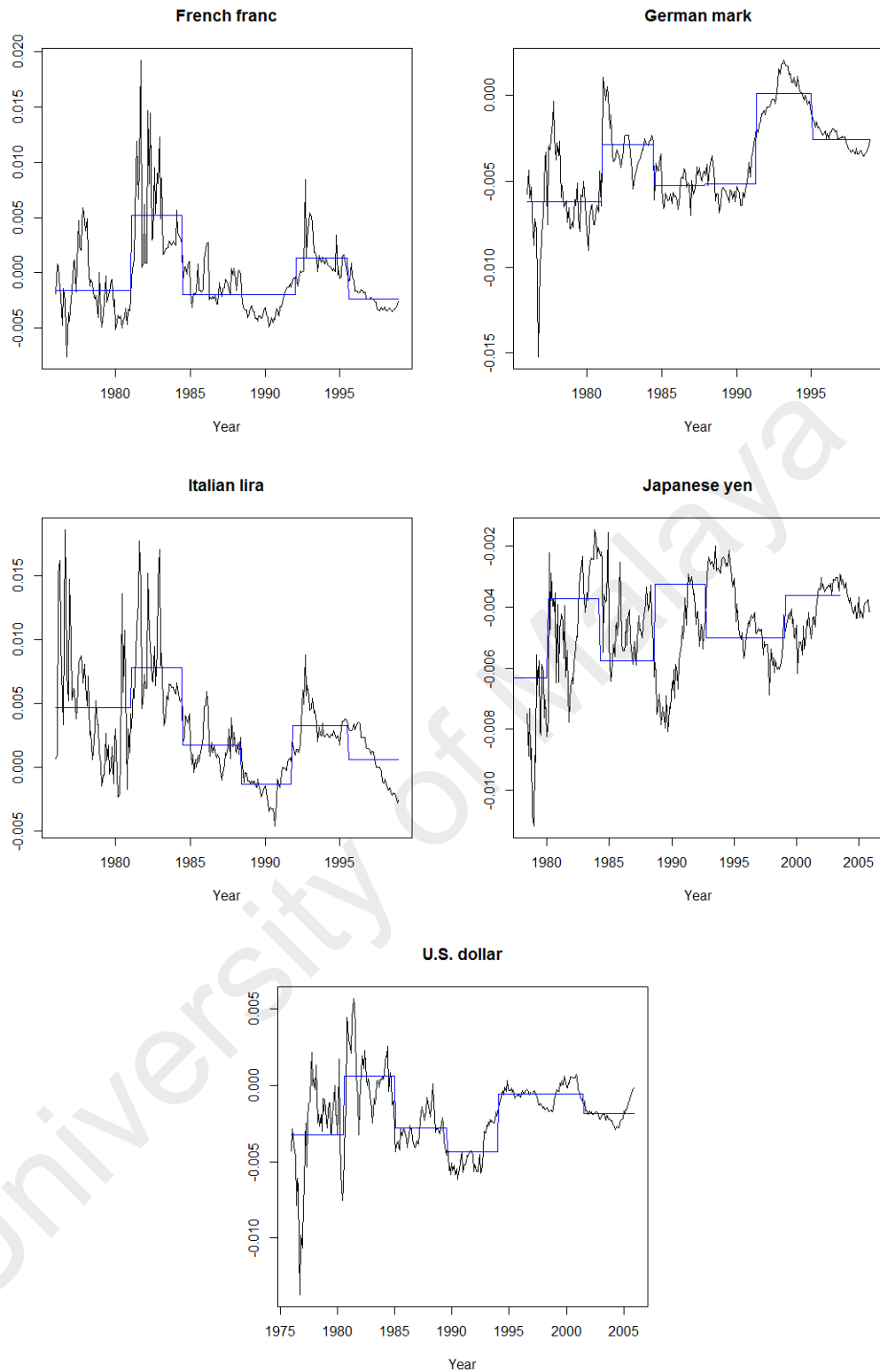


Figure 4.3: Multiple breaks in mean of forward discount³⁶

³⁶ The plot of multiple breaks in mean is using R statistical software code written by Zeileis and Kleiber (2005), available at <http://qed.econ.queensu.ca/jae/2005-v20.5/zeileis-kleiber/>

Table 4.4: Result of structural break in forward discount

	French franc	German mark	Italian lira	Japanese yen	U.S. dollar
$\hat{\delta}_1$	-0.002(0.001)	-0.006(0.0007)	0.005(0.001)	-0.006(0.001)	-0.003(0.001)
$\hat{\delta}_2$	0.005(0.001)	-0.003(0.0006)	0.008(0.001)	-0.004(0.0005)	0.001(0.001)
$\hat{\delta}_3$	-0.002(0.001)	-0.005(0.0003)	0.002(0.0005)	-0.006(0.0007)	-0.003(0.001)
$\hat{\delta}_4$	0.001(0.0004)	-0.005(0.0008)	-0.001(0.0008)	-0.003(0.0005)	-0.004(0.002)
$\hat{\delta}_5$	-0.002(0.0005)	0.0001(0.0006)	0.003(0.0005)	-0.005(0.0002)	-0.001(0.0003)
$\hat{\delta}_6$	-	-0.003(0.0002)	0.001(0.033)	-0.004(0.0001)	-0.002(0.002)
\hat{T}_1	1981:01	1980:12	1981:01	1982:06	1980:07
\hat{T}_2	1984:06	1984:06	1984:06	1986:08	1985:01
\hat{T}_3	1992:01	1987:11	1988:05	1991:01	1989:07
\hat{T}_4	1995:07	1991:04	1991:10	1995:02	1994:01
\hat{T}_5	-	1995:01	1995:07	2001:07	2001:06

1) Asymptotic standard errors are reported in parenthesis.

2) $\hat{\delta}_j$ is the estimated intercept parameter and \hat{T}_j is the estimated break date

4.5.3 Semiparametric Estimates after Break Adjustments

In Table 4.5 and 4.6, we report the long memory estimate of \hat{d}_{MLP} and \hat{d}_{FELW} respectively after adjusting for structural breaks. That is, the estimate of d using the residual series $\hat{u}_t = (f_t - s_t) - \hat{\delta}_j$ where $\hat{\delta}_j$ represents the mean of the forward discount from Bai and Perron's (2003) tests for $j = 1, \dots, m$ breaks from Table 4.4.

Table 4.5: Estimate of \hat{d}_{MLP} with breaks adjustment

	m	\hat{d}_{MLP}	S.E.	90% C.I.
French franc	$T^{0.70}$	0.372	0.105	[0.200, 0.545]
	$T^{0.75}$	0.441	0.086	[0.300, 0.582]
	$T^{0.80}$	0.434	0.078	[0.306, 0.562]
German mark	$T^{0.70}$	0.394	0.094	[0.240, 0.548]
	$T^{0.75}$	0.538	0.084	[0.400, 0.676]
	$T^{0.80}$	0.513	0.070	[0.398, 0.629]
Italian lira	$T^{0.70}$	0.367	0.095	[0.211, 0.523]
	$T^{0.75}$	0.338	0.074	[0.216, 0.459]
	$T^{0.80}$	0.473	0.077	[0.347, 0.599]
Japanese yen	$T^{0.70}$	0.445	0.114	[0.258, 0.633]
	$T^{0.75}$	0.561	0.097	[0.401, 0.721]
	$T^{0.80}$	0.577	0.084	[0.438, 0.715]
U.S. dollar	$T^{0.70}$	0.581	0.074	[0.458, 0.703]
	$T^{0.75}$	0.681	0.067	[0.572, 0.791]
	$T^{0.80}$	0.761	0.060	[0.662, 0.859]

- 1) Selection of m following Choi and Zivot (2007)
- 2) \hat{d}_{MLP} is from Eq. (4.12)
- 3) 90% confidence intervals (CI) are constructed from Eq. (4.13)

They are significant drop in long memory parameter as compare to without breaks adjustment for both semiparametric methods. In contrast to Choi and Zivot (2007), the finding of French franc and German mark after break adjustment are slightly more persistence even though the number breaks for both of these currencies is the same. Choi and Zivot (2007) reported that French Franc and German mark after break

adjustment is 0.295 and 0.229 respectively based on *MLP* at ordinate level of $m = T^{0.70}$. Nonetheless, the result for Italian lira is quite similar to Choi and Zivot (2007). After break adjustment, they reported $\hat{d}_{MLP} = 0.356$ while in this study $\hat{d}_{MLP} = 0.367$ for $m = T^{0.70}$.

Overall, we still found evidences of long memory in all currencies' forward discount even after break adjustment. This finding is in line with Choi and Zivot (2007), where after break adjustment, the long memory process still exists. Thus, based on Choi and Zivot's (2007) method, the structural breaks fails to explain the long memory behaviour found in G7 currencies' forward discount. Nevertheless, we might suspect that long memory estimation result reported in Maynard and Phillips (2001), where for all currencies $0.882 \leq \hat{d}_{MLP} < 1$, will be lower if adjusted for structural breaks.

Table 4.6: Estimate of \hat{d}_{FELW} with breaks adjustment

	m	\hat{d}_{FELW}	S.E.	90% C.I.
French franc	$T^{0.60}$	0.363	0.093	[0.211, 0.515]
	$T^{0.65}$	0.279	0.080	[0.147, 0.411]
	$T^{0.70}$	0.464	0.070	[0.349, 0.579]
German mark	$T^{0.60}$	0.298	0.093	[0.146, 0.450]
	$T^{0.65}$	0.590	0.080	[0.458, 0.722]
	$T^{0.70}$	0.655	0.070	[0.540, 0.770]
Italian lira	$T^{0.60}$	0.341	0.093	[0.189, 0.493]
	$T^{0.65}$	0.359	0.080	[0.227, 0.491]
	$T^{0.70}$	0.388	0.070	[0.273, 0.503]
Japanese yen	$T^{0.60}$	0.439	0.088	[0.295, 0.583]
	$T^{0.65}$	0.635	0.076	[0.510, 0.760]
	$T^{0.70}$	0.613	0.066	[0.505, 0.721]
U.S. dollar	$T^{0.60}$	0.327	0.086	[0.186, 0.468]
	$T^{0.65}$	0.545	0.074	[0.424, 0.666]
	$T^{0.70}$	0.627	0.064	[0.522, 0.732]

1) Selection of m as suggested by Hassler and Meller (2014)

2) \hat{d}_{FELW} is from Eq. (4.20)

3) 90% confidence intervals (CI) are constructed from Eq. (4.21)

4.5.4 Spurious Long Memory Tests Result³⁷

In Table 4.7 below we report the estimate of \hat{d} and \bar{d} , the value of W_c and $\hat{\eta}_\mu$ (KPSS) of forward discount. As shown in Shimotsu (2006), his test of long memory versus structural break performs well even with a small sample size of $T = 240$ observations.

Table 4.7: Shimotsu's (2006) 'simple but effective' long memory test result

	$m = T^\alpha$	\hat{d}	\bar{d}		W_c		$\hat{\eta}_\mu$
			$b = 2$	$b = 4$	$b = 2$	$b = 4$	
Canadian dollar	20	0.206	0.441	0.543	1.809	2.103	0.109
	40	0.611	0.677	0.887	0.021	4.754	0.030
	60	0.758	0.819	1.007	0.037	11.004*	0.021
French franc	20	0.705	0.833	0.911	0.716	14.019*	0.087
	40	0.667	0.843	0.948	3.511	6.430	0.095
	60	0.807	0.848	0.820	0.041	3.463	0.062
German mark	20	0.703	0.885	1.192	9.268*	5.018	0.066
	40	0.937	1.072	0.930	2.863	2.263	0.030
	60	0.917	1.042	1.043	3.742	4.804	0.032
Italian lira	20	0.572	0.830	1.302	3.483	1.171	0.219
	40	0.536	0.802	0.900	9.733*	3.075	0.248
	60	0.550	0.794	0.855	15.815*	8.543*	0.235
Japanese yen	20	0.793	0.879	0.990	0.113	2.994	0.070
	40	0.798	0.981	1.107	2.891	3.873	0.068
	60	0.765	0.898	1.003	3.917*	4.147	0.078
U.S. dollar	20	0.617	0.712	1.000	3.458	2.502	0.059
	40	0.737	1.046	1.134	8.401*	9.887*	0.050
	60	0.760	0.984	1.044	7.841*	10.065*	0.045

1) * indicates rejection of the null at 5% level. $\chi_{0.95}^2(1) = 3.84, \chi_{0.95}^2(3) = 7.82$

2) \hat{d} is estimated using Eq. (4.20)

3) W_c is from Eq. (4.25)

4) $\hat{\eta}_\mu$ is statistic of $\log|v_i|$ of Eq. (4.27)

This supports the use of this test in this study, given that our sample of data are consider small in the long memory literature. However, we may found some local variation in estimation of d due to small periodogram ordinates of m . The sample split is set to $b = \{1,2,4\}$ and the periodogram ordinates of $m = \{20,40,60\}$.³⁸ The selection of m and T must fulfill the requirement that T/b and m/b are integer. Due to that fact, sample adjustment is needed for Japanese yen where the sample is set from June 1978

³⁷ Shimotsu spurious long memory test is conducted using MATLAB code written by Katsumi Shimotsu which is available at http://shimotsu.web.fc2.com/Site/Matlab_Codes.html. The Qu (2011) test is conducted using code written in R statistical software by Zhongjun Qu and available at <http://people.bu.edu/qu/code.htm>

³⁸ $b = 1$ means no sample splits, which actually reported in column 3 of Table 4.7

to January 2005, where $T = 330$ observations. No sample adjustment is needed for the rest of currencies in order to conduct the test.

For all currencies, the estimates of long memory parameter without sample splitting are in the range of $0.5 < \hat{d} < 1$, a nonstationary long memory process with exception to Canadian dollar for $m = 20$. The W_c test of null constancy in d is rejected only once for Canadian dollar, French franc, German mark and Japanese yen even though we see some local variation in estimation of d . Together with $\hat{\eta}_\mu$ test statistics of d th differencing, these four currencies shows evidence against spurious long memory process. However, for Italian lira and U.S. dollar, the findings are rather mix where evidence non-constancy is rejected at multiple level of m while the $\hat{\eta}_\mu$ statistics do not reject the null of $I(\hat{d})$ in all cases. Overall, the findings do not find strong evidence in favour of a spurious long memory process of forward discount for all currencies.

Table 4.8: Qu's (2011) spurious long memory test

	$m = T^{0.55}$	$m = T^{0.60}$	$m = T^{0.65}$	$m = T^{0.70}$
Canadian dollar	0.814	0.888	1.088	1.118
France franc	0.477	0.457	0.291	0.941
German mark	0.686	0.673	0.818	0.699
Italian lira	0.551	0.469	0.694	0.659
Japanese yen	0.939	0.890	0.888	0.766
U.S. dollar	0.596	0.847	0.711	0.842

1) The trimming proportion is set to $\varepsilon = 0.05$ as suggested by Qu (2011) for small sample.

2) * denotes rejection at 5% significant level

3) Asymptotic critical value for $\varepsilon = 0.05$ at 5% and 1% are 1.155 and 1.426 respectively.

The Qu's (2011) spurious long memory test is reported in Table 4.8, where the test is conducted with $\varepsilon = 0.05$ as trimming parameter due to small sample. To reduce the impact of short-memory component in small sample, we apply the filtering procedure as suggested by Qu (2011). The test fail to reject the null of stationary long memory at all level of periodogram ordinates in all currencies involved. The findings

above show strong indication that forward discount for all G7 currencies' are stationary long memory process rejecting the alternative that is affected by regime change or smoothly varying trend. Given the findings of Shimotsu's (2006) and Qu's (2011) tests, we may conclude that G7 currencies' forward discounts are a true long memory process.

4.6 Conclusion

This chapter set out to determine whether the findings of long memory properties in previous studies are true or spurious. By using monthly forward discount series of G7 countries, the approaches of this chapter are two folds. Firstly, we apply Choi and Zivot's (2007) structural breaks adjustment method where we estimate the long memory parameter twice; before and after structural breaks adjustment. We extend Choi and Zivot's (2007) approach of long memory estimation by using two-step feasible exact local Whittle (FELW) estimator of Shimotsu and Phillips (2006). The use of FELW beside modified log-periodogram (MLP) of Phillips (2007) and Kim and Phillips (2000, 2006) are due to two reasons; 1) FELW is more efficient asymptotically and 2) FELW is robust with respect to heteroskedasticity of a certain degree (Robinson and Henry, 1999; Shao and Wu, 2007). Secondly, we apply statistical test of spurious long memory by Shimotsu (2006) and Qu (2011) on the series.

In the first approach, the results of MLP and FELW found evidence of long memory before and after structural breaks adjustment, even though we found evidence of structural breaks in most currencies based on Bai and Perron's (2003) model. The estimates of long memory parameter are significantly less after structural breaks adjustment. However, structural breaks fail to explain the long memory behaviour found in G7 currencies' forward discount.

Based on second approach, the result of Shimotsu's (2006) long memory test findings do not found strong evidence in favour of spurious long memory process of forward discount for all currencies. This further supported by Qu's (2011) spurious long memory test result, where the test fail to reject the null of stationary long memory at various level of periodogram ordinates m . Overall, the results from structural break adjustment and statistical test of spurious long memory are in favour of true long memory process of G7 currencies used in this study.

University of Malaya

Chapter 5: Nonlinearity in Forward Discount

5.1 Introduction

In this chapter, the discussion aims to answer the second objective of this study, specifically nonlinearity of forward discount.³⁹ Based on this objective, the presence of nonlinearity will be determined statistically followed by examination of the existence of ‘band inaction’ based on transaction cost. Finally, the forecasting performance of nonlinear versus linear models is conducted.

5.2 Stochastic Process of Forward Discount

The empirical evidence of stationarity in forward discount is decidedly mixed. In previous literatures, findings based on various currencies, time periods and frequencies are inconclusive as discussed in Engel (1996).⁴⁰ However, nonstationarity of forward discount is not seemed particularly plausible in an economic sense.

In previous studies, empirical findings uncovered cointegration of nominal spot exchange rates (e.g. Alexander and Johnson, 1992; Baillie and Bollerslev, 1989, 1994a) which seem to contradict the market efficiency hypothesis. This is because a cointegrated system of nominal spot exchange rates implies the presence of predictability of returns in at least one currency. However, Crowder (1994) argues that the cointegrating relationship of nominal spot exchange rates may reflect a time-varying

³⁹ Refer to section for 1.4 detail discussion.

⁴⁰ Earlier findings of non-stationary are Crowder (1992,1994), Lewis (1995), Evans and Lewis (1993,1995) and Luintel and Paudyal (1998). Evidence of stationarity in forward discount are Mark et al. (1997), Baillie and Bollerslev (1989), Barkoulas et al. (2003), Clarida and Taylor (1997), Horvath and Watson (1995) and Meese and Singleton (1982).

risk premium evident in several currencies' returns. With the assumption of risk aversion, efficiency of foreign exchange market implies that a time-varying risk premium and the error-correction term from the above-cointegrated system must be covariance stationary. However, since risk premium is not directly observable, it shares the same time series properties as forward discount as shown in Eq. (2.18). If one finds that forward discount is nonstationary, similar property will also exist in risk premium. Nevertheless, the cointegrating relationship is stationary by definition (see Engel and Granger, 1987), as such empirical findings of nonstationarity in forward discount is implausible.

The theory of covered interest parity (CIP) stated that when it holds, the forward discount is simply the difference between the domestic and foreign interest rate as shown in Eq. (2.4). Since interest rate is bounded below by zero, the linear combination of this rate cannot have unbounded mean, which is a true unit root process. Regardless of whether CIP holds, empirical findings nonstationarity in forward discount in previous studies findings may suggest theoretical failure of CIP model. This is difficult to comprehend given that CIP model is one of the building block in the area of International Finance.

Nonstationarity of forward discount will invalidate the standard statistical inference of regression-based test in which FRUH testing is based on. This will lead the regression-based test of Eq. (2.7) to become unbalanced. Thus, anomalous findings in FRUH may partly be explained due to the finding of unit root in the forward discount (see Baillie and Bollerslev, 2000). This highlights the importance of determining the stationarity of forward discount given that from econometric perspective, nonstationarity of forward discount is the possible cause of forward bias puzzle.

Various unit root test procedures have been applied in testing forward discount. It is a well-known fact that unit root test is not good at distinguishing a series with

characteristic root that is close to unity⁴¹ and structural change. In the case of structural change, the standard Dickey-Fuller test is biased towards the non-rejection of a unit root. However, allowing for multiple breaks blurs the distinction between a unit root process and stationary series with breaks (Hansen, 2001) and the actual test creates difficulties of practical implementation (Perron, 2006).

Another recent strand of research focuses on potential issues arises due to nonlinearities in the underlying data generating process (DGP) have on unit root testing. The idea of nonlinearities in macroeconomic variable can be traced back in a research conducted by Stock and Watson (1996), where there are widespread of parameter instability in univariate and bivariate autoregressive models of large set of macroeconomic data set. As argued by Enders and Granger (1998), if nonlinearities are prevalent under the alternative of stationarity, a linear unit root test suffers from lack of power. Taylor et al. (2001) also support this view where he argues; if the adjustment is nonlinear, linear test will fail to support mean reversion if the process spends considerable time inside the unit root band.

5.3 Nonlinearities

Previous empirical studies in analysing the stationarity of forward discount rely on linear frameworks. There are several reasons that we might suspect that forward discount is a nonlinear process. One of the reasons is the existence of transaction cost. This idea is pioneered by Dumas (1992), where he developed a general equilibrium model of exchange rate determination in spatially separated markets with significant cost of international trade. Based on the model, the transaction cost will create a 'band of inaction', where inside the band; there is no adjustment in deviation from equilibrium

⁴¹ This is exactly the case of forward discount

that takes place. However, outside the band, the process becomes mean reverting since the benefits of arbitrage exceed the cost. Taylor (2001) also supports the existence of a 'band of inaction' in his seminal paper. He argued that fixed and variable trading costs result in non-action of market participants when there is no arbitrage. The mean reversion only occurs when the prices move a lot.

In another study by Obstfeld and Taylor (1997), they argue the existence of two regimes due to transaction cost. One show slow or non-convergence when the price differences in two countries are small and the other show rapid convergence when the price difference exceeds transaction cost. It should be noted all the studies above suggest discrete changes from one regime to another. Some studies are in favour of smooth threshold models (e.g. Taylor and Peel, 2000; Killian and Taylor, 2003). However, one should expect that the market participant would react quickly when opportunity such as arbitrage exist causing an abrupt pattern rather than smooth.

The idea of limits to speculation hypothesis of Lyons (2001) may also explain nonlinearities in forward discount. The model emphasizes on the importance of Sharpe Ratio in determining the investment strategies.⁴² It states that when the Sharpe Ratio is higher than a threshold level, the deviation from uncovered interest parity (UIP) will be high enough to be viewed as arbitrage opportunity. The implication of the model is that it creates a band of Sharpe Ratio, which results in a band of forward discount where UIP does not hold.⁴³ Other possible explanations of nonlinearities of forward discount are brought forward by Sarno et al. (2004), which are automatic trading rules, heterogeneous belief and tendency of traders to wait for large arbitrage opportunities before entering the market.

⁴² Sharpe Ratio is defined as $\frac{E[R_s - R_{rf}]}{\sigma_s}$, where $E[R_s]$ is the expected return on the strategy, R_{rf} is the risk-free interest rate and σ_s is the standard deviation of the returns of the strategy (see Sharpe (1985)).

⁴³ Some researchers used UIP testing instead of FRUH testing. The words are used interchangeably in the research.

Even though evidence of nonlinearities in foreign exchange rate has been found by previous studies,⁴⁴ it does not distinguish between nonlinearity from unit root behaviour in the exchange rate. One important aspect is that if prior assumptions of stationarity are not valid and the variables have unit root, it will lead to incorrect inferences of the test of linearity versus threshold alternatives. This is due to the nonstandard asymptotic distribution of the test.⁴⁵ More than a decade ago, Caner and Hansen (2001), afterwards CH, has developed a symmetric threshold autoregressive (TAR) model with an autoregressive unit root. CH is the first to combine the presence of unit root type of nonstationarities and threshold type of nonlinear dynamics. Their major contribution was the development of a new asymptotic theory for detecting the presence of threshold effects in a series, which was restricted to be a unit root process under the null of linearity. The model allows for inner no-arbitrage band with small disequilibria while capturing mean reversion to shocks outside the no-arbitrage band. In this test, we allowed for general autoregressive orders and no restriction in the coefficients across regime (Basci and Caner, 2005).

5.4 Methodologies

5.4.1 Nonlinear Unit Root Test Model

The model is based on threshold autoregression (TAR) of the following:

$$\Delta fd_t = \theta'_1 x_{t-1} 1_{\{Z_{t-1} < \lambda\}} + \theta'_2 x_{t-1} 1_{\{Z_{t-1} \geq \lambda\}} + e_t, \quad t = 1, 2, \dots, T, \quad (5.1)$$

where fd_t represents forward discount, $x_{t-1} = (fd_{t-1}, 1, \Delta fd_{t-1}, \dots, \Delta fd_{t-k})'$, $1_{\{\cdot\}}$ is an indicator function and e_t is i.i.d. error. The threshold variable, $Z_{t-1} = fd_{t-1} - fd_{t-m-1}$, with m represents delay parameter and $1 \leq m \leq k$. The λ represents the threshold value and is unknown, and takes on value in the compact interval $\Lambda =$

⁴⁴ See Baum et al. (2001), Clarida et al. (2003), Kilian and Taylor (2003), Panos et al. (1997), Obstfeld and Taylor (1997), Sarantis (1999) and Taylor et al. (2001).

⁴⁵ See Bec et al. (2002)

$[\lambda_1, \lambda_2]$ where λ_1 and λ_2 are picked so that $P(Z_{t-1} \leq \lambda_1) = .15$ and $P(Z_{t-1} \leq \lambda_2) = .85$. The specification of the threshold variables in difference rather than level is necessary due to the econometric theory developed does not allow levels form (Caner and Hansen, 2001).

The components of θ_1 and θ_2 can be separated into:

$$\theta_1 = \begin{pmatrix} \rho_1 \\ \beta_1 \\ \alpha_1 \end{pmatrix}, \quad \theta_2 = \begin{pmatrix} \rho_2 \\ \beta_2 \\ \alpha_2 \end{pmatrix} \quad (5.2)$$

where (ρ_1, ρ_2) are the slope on fd_{t-1} , (β_1, β_2) represents the intercept coefficients and (α_1, α_2) are the slope on dynamic regressors of $(\Delta fd_{t-1} \cdots \Delta fd_{t-k})$ in the two regimes.

The model in Eq. (5.1) is estimated by least square (LS) of;

$$\Delta fd_t = \hat{\theta}'_1(\lambda)x_{t-1}1_{\{Z_{t-1} < \hat{\lambda}\}} + \hat{\theta}'_2(\lambda)x_{t-1}1_{\{Z_{t-1} \geq \hat{\lambda}\}} + \hat{e}_t(\lambda) \quad (5.3)$$

where $\hat{e}_t(\lambda)$ is the LS residuals and $\hat{\sigma}^2(\lambda) = T^{-1} \sum_{t=1}^T \hat{e}_t(\lambda)^2$ denotes the residuals variance from the LS estimation. The threshold parameter is estimated by minimizing $\hat{\sigma}^2(\lambda)$:

$$\hat{\lambda} = \arg \min_{\lambda \in \Lambda} \hat{\sigma}^2(\lambda) \quad (5.4)$$

The above optimal threshold $\hat{\lambda}$ is then plugged into Eq. (5.3) to determine other parameters of interest. The first test statistic is to determine the existence of nonlinearity in the series due to threshold, where the threshold effects disappear under the joint hypothesis of:

$$H_0: \theta_1 = \theta_2 \quad (5.5)$$

The test statistics is :

$$\sup_{\lambda \in \Lambda} W_t(\lambda) = \sup_{\lambda \in \Lambda} T \left(\frac{\hat{\sigma}_0^2}{\hat{\sigma}^2}(\hat{\lambda}) - 1 \right) \quad (5.6)$$

where $\hat{\sigma}^2$ and $\hat{\sigma}_0^2$ are the residuals variances from the TAR and linear autoregressive (AR) models respectively.

Based on model (5.1), parameter (ρ_1, ρ_2) determines the stationarity of the process, which result in three different hypothesis of:

$$H_0 : \rho_1 = \rho_2 = 0 \quad (5.7)$$

a unit process of forward discount,

$$H_1 : \rho_1 < 0 \text{ and } \rho_2 < 0 \quad (5.8)$$

Two-regime stationary threshold autoregressive (AR) process and

$$H_2 : \begin{cases} \rho_1 < 0 \text{ and } \rho_2 = 0 \\ \text{or} \\ \rho_1 = 0 \text{ and } \rho_2 < 0 \end{cases} \quad (5.9)$$

a partial unit root case. If the H_2 hypothesis holds, the process of fd_t behaves like a unit root process in one regime, yet behaves like stationary process in the other regime. Thus, the process under H_2 process is nonstationary, but it is not the classical unit root process (Caner and Hansen, 2001). Since H_1 hypothesis is one-sided, the Wald's one-sided test statistic of $H_0: \rho_1 = \rho_2 = 0$ vs H_1 is:

$$R_{1T} = t_1^2 1_{\{\hat{\rho}_1 < 0\}} + t_2^2 1_{\{\hat{\rho}_2 < 0\}} \quad (5.10)$$

where t_1 and t_2 are the t -ratios of the $\hat{\rho}_1$ and $\hat{\rho}_2$ respectively of Eq. (5.3). However, in order to discriminate between H_1 and H_2 , we cannot rely on R_{1T} statistics above (Caner and Hansen, 2001). As suggested by Caner and Hansen (2001), in order to test H_0 vs H_2 , we rely on the negative of the $-t_1$ and $-t_2$ statistics.

5.5 Unit Root Test

The Table 5.1 below shows the results of unit root test based on ADF, where the number of lag is determined by modified Akaike information criterion (MAIC) proposed by Ng and Perron (2001) with the maximum bounds of 12 lags is set for each currency's forward discount. This finding provides the link with previous studies and the non-linear unit root test used in this chapter. The test is conducted with intercepts for all currencies. The nullity of unit root is rejected for Canadian dollar, French franc, Japanese yen and U.S. dollar, but not for German mark and Italian lira. The result for German mark is similar to Crowder (1994) and Choi and Zivot (2007). However, in Crowder (1994) and Choi and Zivot (2007), they fail to reject the null of unit root for Canadian dollar, which in contrast with our finding. For French franc, our result is similar to Choi and Zivot (2007), rejecting the null of unit root. However, in Choi and Zivot (2007), the unit root test is conducted using ADF-GLS t -statistics of Elliot et al. (1996). Furthermore, Crowder (1994) sample period is 1974:01 to 1991:12 and in this study; it is from 1976:01 to 1998:12. As noted in a survey by Engle (1996), there are mixed findings regarding stationarity of forward discount that cover different currencies and time periods.

The inability to reject the null hypothesis should not be regarded as *prima facie* evidence against mean-reverting behaviour. It is widely known that the ADF test has low power against a stationary alternative with a high level of persistency where the root of the autoregressive (AR) process close to unity or if the data-generating process (DGP) is nonlinear (Caner and Hansen, 2001). In the context of forward discount, Sakoulis et al. (2010) showed that the forward discount is indeed a highly persistent

$AR(1)$ process. Nonetheless, our finding based on unit root test adds further confusion on the mix finding of stationarity issue of forward discount, as highlighted in Engle (1996).

Table 5.1: ADF unit root test

	Canadian dollar	French franc	German mark	Italian lira	Japanese yen	U.S. dollar
Lag	10	9	10	10	12	12
ADF	-4.186 (0.001)	-3.393 (0.012)	-2.393 (0.145)	-1.940 (0.314)	-3.845 (0.003)	-3.533 (0.008)
ρ_1	-0.141 [0.034]	-0.151 [0.044]	-0.058 [0.024]	-0.073 [0.038]	-0.123 [0.032]	-0.093 [0.026]

1) Standard error and p -value are in prathensis and brackets respectively.

2) Lags are determined by MAIC with maximum upper bound of 12 lags.

5.6 Findings

5.6.1 Non-Linearity Test⁴⁶

Table 5.2 reports the result of the Wald test of Eq. (5.6). With the null of linearity, signification rejection of the test suggest the existence of threshold hence nonlinearity of forward discount. The reported critical values are based on bootstrap (BCV) of 5,000 replications corresponding to the optimal delay parameter m . The optimal delay parameters are determined endogenously, where it minimizes the error sums of squares of the TAR model.

As reported in Table 5.2, the results show strong evidence for threshold effect hence nonlinearity at 1% level in all currencies. With the delay parameters represent the change in forward discount in months, larger value of m suggest the longer it takes for

⁴⁶ The linearty test and the following TAR unit root test is conducted using GAUSS code written by Bruce Hansen and available at http://www.ssc.wisc.edu/~bhansen/progs/progs_threshold.html

market participants to react to deviations from the relationship that link spot and forward rates. This suggests that arbitrage opportunities are exploited rapidly for countries with low delay parameter.

To give an example in the case of Japanese yen, the forward discount switches regime based on the value of the forward discount 1 month earlier implying the swift exploitation of arbitrage opportunities for this currency. However, it is not clear why currencies like German mark and Canadian dollar have high delay parameter since we expect the delay to be small for deviations begin to adjust in response to a shock. Nonetheless, with these evidences, it clearly implies misspecification of the functional form in the regression for the ADF test used in Table 5.1 since the construction of ADF test assumed linearity, which also used in Crowder (1994) and Choi and Zivot (2007). In the next section will apply TAR unit root test for all currencies based on their delay parameter.

Table 5.2: Non-linearity test result

	m	Wald	10% BCV	5% BCV	1% BCV
Canadian dollar	7	44.9	28.1	31.3	37.1
French franc	4	102.0	29.5	33.5	43.1
German mark	6	55.8	30.2	33.9	42.5
Italian lira	2	73.4	31.0	35.4	46.1
Japanese yen	1	77.5	32.7	36.8	45.4
U.S. dollar	5	83.0	34.3	39.0	49.5

1) m is the optimal delay parameter

2) Bootstrap critical value (BCV) is based on 5,000 replications

5.6.2 TAR Unit Root Test

In Table 5.3, we report the TAR unit root test based on the delay parameter m selected by Wald test earlier. Based from the result in the previous section, we confirm the existence of two-regime symmetric model of TAR where the parameter of interest

represented in Eq. (5.2). The process switch from first regime (inside band) to second regime (outside band) based on threshold/band in column 6 of Table 5.3. To give an example in the case of Japanese yen, if the change in the forward discount in the past 1 month (*i. e.* $(fd_{t-1} - fd_{t-m-1})$ where it is $m = 1$ for Japanese yen) are above 0.0005, the process will switch to the second regime (outside band).

The stationarity of the two regimes are determined based on parameter (ρ_1, ρ_2) of Eq. 5.2. It should be noted that parameter ρ_1 represents the first regime (inside the band) while parameter ρ_2 represents the second regime (outside the band). To determine whether the two regimes are unit root process versus both are stationary regimes (H_0 vs H_1), we relied on R_{1T} test statistic results. If significant, both regimes have unit root hypothesis are rejected. Based on R_{1T} results, Canadian dollar, French franc, Japanese yen and U.S. dollar reject the null (H_0) that both regimes have unit root. However, for German mark and Italian lira, the H_0 fails to be rejected suggesting that both regimes of these currencies have unit root.

In order to discriminate further between unit root processes versus partial unit root process (H_0 vs H_2), we cannot rely on R_{1T} (Caner and Hansen, 2001). For that purpose, we used t_1 and t_2 that represent the first regime (inside band) and the second regime (outside band) respectively. Based on the results reported in column 4 and 5 of Table 5.3, all forward discount's inside band fail to reject the null of $\rho_1 = 0$ (based on t_1), or inside band has unit root. Looking at t_2 , the null of $\rho_2 = 0$ is rejected for all currencies except for German mark. Thus for German mark, the inside and outside band are unit root process.

Overall, we have clear evidences that suggest Canadian dollar, French franc, Japanese yen and U.S. dollar's forward discount are a partial unit root process that consists of one nonstationary and one stationary regime. In the inside band, these

currencies' forward discounts follow a unit root behaviour suggesting inaction of market participants due to transaction cost. As such, in inside band, no adjustment in deviation from equilibrium takes place where the deviation follows a unit root process. In the outside band, the benefit of arbitrage exceeds the cost, where the process switches abruptly to become mean reverting towards the transaction cost band.

Looking closely at German mark and Italian lira results, the R_{1T} test fail to reject H_0 of unit root process in both regime due to the fact the nonstationarity of first regime is dominant, where both currencies has 84% of the observation in the first regime. Furthermore, for Italian lira, the t_2 for outside band reject the null of unit root. Thus, for Italian lira, the result is in favour of partial unit root process despite failure to reject the of unit root process in both regime based on R_{1T} test.

Overall, the results are in favour of partial unit root process of forward discount. This explains the mixed finding of stationarity in previous studies, for example, earlier evidences of nonstationary are Crowder (1992,1994), Lewis (1995), Evans and Lewis (1993,1995) and Luintel and Paudyal (1998). On the other hand, earlier findings of stationarity in forward discount are Mark et al. (1997), Baillie and Bollerslev (1989), Barkoulas et al. (2003), Clarida and Taylor (1997), Horvath and Watson (1995) and Meese and Singleton (1982)). Forward discount is shown to be globally stationary albeit high persistence since most observations lies inside the dominant unit root regime. The ADF unit root test will fail to reject the null of unit root since $I(1)$ part of the series will dominate asymptotically (see Chong, 2001; Taylor, 2005). The first regime is dominant in all countries where it fluctuates between 61% and 85% of the observations in the countries analysed.

Table 5.3: TAR unit root test

	m	R_{1T}	t_1	t_2	λ	%		ρ_1	ρ_2
		Unit root test	Inside band	Outside band	Threshold/band	Inside band	Outside band	Inside band	Outside band
Canadian dollar	7	27.6***	1.49	5.04***	0.0012	0.61	0.39	-0.079(0.053)	-0.215(0.043)
French franc	4	56.7***	1.15	7.44***	0.0023	0.70	0.30	-0.067(0.058)	-0.527(0.071)
German mark	6	3.27	1.72	0.55	0.0023	0.84	0.16	-0.041(0.024)	-0.046(0.084)
Italian lira	2	9.33	1.63	2.58*	0.0029	0.84	0.16	-0.071(0.044)	-0.208(0.080)
Japanese yen	1	14.0**	2.36	2.91**	0.0005	0.71	0.29	-0.090(0.038)	-0.156(0.054)
U.S. dollar	5	12.6**	1.42	3.26**	0.0021	0.85	0.15	-0.041(0.029)	-0.157(0.048)

1) m is the delay parameter

2) Figures in parentheses are standard errors.

3) ***, ** and * denotes significant at the 1%, 5% and 10% respectively.

4) λ is from Eq. (5.4)

5) R_{1T} is from Eq. (5.10)

5.6.3 Out of Sample Forecast⁴⁷

In this section, we compare the forecasting performance of linear $AR(k)$ model against threshold model of Eq. (5.1) for forward discount. All original samples are cut into two parts. Beginning from the end of the first part, we obtain 1,3,6,9 and 12 months ahead forecast of forward discount for both model of linear $AR(k)$ and non-linear TAR model. For Canadian dollar and U.S. dollar, the first part consists of monthly observation from 1976:01-2004:12. The first part of French franc, German mark and Italian lira consist of monthly observation from 1978:06-2004:12 while for Japanese yen, 1978:06-2004:12.

The computed forecast error is based on root mean squared error (RMSE) statistic. It is a naïve in the sense that the threshold and TAR parameters in sample are kept fixed, where only the data points are updated after each forecast. In general, the forecast of nonlinear TAR model outperform linear $AR(k)$ model for Canadian dollar, French franc, Italian lira and Japanese yen regardless of forecast period as reported in Table 5.4. For German mark, both models are indifferent except for 9 months ahead forecast where linear $AR(k)$ model outperform nonlinear TAR model. However, for U.S. dollar, overwhelming evidence shows that linear model performs better in forecasting exercise. The performance of both model in forecasting are summarized in Table 5.5 where the ratio is derived by dividing the RMSE of linear model by RMSE of TAR model. For ratio greater than unity, it shows that TAR model perform well compare to linear model in forecasting exercise.

⁴⁷ The forecast for linear $AR(k)$ model versus TAR model is using GAUSS code written by Erdem Basci and Mehmet Caner. The GAUSS code are available at:
<http://www.degruyter.com/view/j/snde.2005.9.4/snde.2005.9.4.1273/snde.2005.9.4.1273.xml?format=INT>

Table 5.4: Out of sample forecasting: Linear $AR(k)$ versus nonlinear TAR model

Period	Linear: RMSE					TAR:RMSE				
	1	3	6	9	12	1	3	6	9	12
Canadian dollar	0.0004	0.0004	0.0007	0.0007	0.0006	0.0002	0.0002	0.0004	0.0004	0.0006
French franc	0.0006	0.0007	0.0010	0.0014	0.0016	0.0004	0.0004	0.0006	0.0007	0.0011
German mark	0.0001	0.0002	0.0002	0.0002	0.0002	0.0001	0.0002	0.0002	0.0003	0.0015
Italian lira	0.0009	0.0009	0.0015	0.0018	0.0021	0.0006	0.0007	0.0015	0.0016	0.0015
Japanese yen	0.0002	0.0003	0.0003	0.0004	0.0005	0.0001	0.0003	0.0003	0.0003	0.0023
U.S. dollar	0.0001	0.0002	0.0003	0.0006	0.0009	0.0001	0.0004	0.0006	0.0009	0.0013

- 1) RMSE stands for root mean squared error
- 2) The forecast begin after the last period of 2004:12 for Canada, Japan and US. For France, Germany and Italy the forecast begin after 1997:12. The period represents month. For example, period 3 Italy represent forecast error made for 1998:03.

Table 5.5: Ratio

Period	1	3	6	9	12
Canadian dollar	2	2	1.75	1.75	1
French franc	1.5	1.75	1.67	2	1.45
German mark	1	1	1	0.67	0.13
Italian lira	1.5	1.29	1	1.06	1.4
Japanese yen	2	1	1	1.33	0.22
U.S. dollar	1	0.5	0.5	0.67	0.69

- 1) Ratio represents the RMSE of linear model divided by RMSE of TAR. Figure greater than unity shows the superior performance of TAR to linear model

5.7 Conclusion

This chapter has provided a different approach in understanding the issue of mixed finding of stationarity in forward discount as reported in previous studies. The present study is designed to determine the stationarity of forward discount if we allow the series to be nonlinear, where the nonlinearity and nonstationarity is test simultaneously. The nonlinearity in forward discount is attributed to, inter alia, transaction cost. In standard unit root tests, the process are assume linear in such a way that the null of a linear unit root is tested against a linear stationary model.

We found strong evidence of nonlinearity in all currencies. Next, two unit root tests, one-sided Wald test and t-ratio tests are used to currencies that show evidence of

nonlinearity. The results show that forward discount has two regimes threshold model in which it display unit root behaviour inside a band and become mean reverting outside the band. This finding support the idea that transaction cost creates a 'band of inaction' when there is no arbitrage. With no arbitrage, no adjustment in deviation from equilibrium takes place in this regime resulting unit root behaviour in this regime. The process than become mean reversion as the benefit of arbitrage exceeds the cost

Most of the observations lie inside the band, which might explain unit root finding in previous studies. As for this chapter, the German mark and Italian lira fail to reject the null of unit root based on ADF test since 85% of their observations are inside the unit root regime. As such, forward discount is globally stationary albeit high persistency since most observations lies inside the dominant unit root regime.

The forecasting performance of TAR model outperform linear (*AR*) model in most currencies involved. The implication of this chapter's finding suggests that nominal interest rate differential between domestic and foreign country as shown in Eq. (2.4) will be a partial unit root process. Furthermore, this finding also has major implication towards the standard regression test of Fama (1994) where nonlinearity in forward discount suggest that earlier testing is miss-specified.

Chapter 6: Autoregressive Root Near Unity Process Of Forward Discount

6.1 Introduction

This chapter aims to answer the third objective of this study, specifically root near unity of forward discount.⁴⁸ Based on this objective, the persistency of forward discount based on root near unity *AR* model is investigated using confidence interval approach. Lastly, we investigate whether root near unity process hold throughout the sample period using persistency change test.

6.2 Root near Unity and Confidence Interval

The extent and nature of bias in FRUH testing depends on the type of persistence displayed by the forward discount (Maynard, 2006). There are two forms of persistence: long memory and autoregressive root near unity.⁴⁹ Recently, evidence of structural breaks has been found in the series (Sakoulis et al., 2010). Sakoulis et al. (2010) showed that the forward discount is less persistent after allowing multiple breaks in the forward discount mean. By modelling forward discount as autoregressive one ($AR(1)$) process of G7 countries' currencies, they argued that breaks in mean have exaggerated the persistence in the forward discount.

⁴⁸ Refer to section for 1.4 detail discussion.

⁴⁹ Maynard and Phillips (2001) and Baillie and Bollerslev (1994b) argued that the forward discount has a long memory. Crowder (1994, 1995) and Baillie and Bollerslev (1994a) showed that the forward discount has a root near unity.

In study by Sakoulis et al. (2010), the OLS estimate of $AR(1)$ specification of forward discount is greater than 0.9 for German mark, British pound and Japanese yen before incorporating structural breaks in mean of the series. This finding is expected where it has long be noted in studies by Crowder (1994, 1995) and Evans and Lewis (1995) that the estimation of autoregressive (AR) coefficients on forward discount are generally large, to the extent that unit roots may sometimes be difficult to reject. This high persistency of forward discount has even persuaded Evan and Lewis (1995) and Crowder (1994) that forward discount may contain nonstationary components.

This large estimation or near-unit root problem results in some studies questioning the validity of empirical inference procedures underlying the forward bias puzzle, suggesting potentially serious bias and/or size distortion (e.g. Bekaert and Hodrick, 2001; Maynard, 2003; Roll and Yan, 2000; Tauchen, 2001; Goodhart et al., 1997; Newbold et al., 1998). In a study by Mankiw and Shapiro (1986), they conducted a simulation based on Eq. (2.7) for a sample size of 200 where the autoregressive root in the regressor is close to but strictly less than 1. They showed that actual rejection rates for a nominal 5% test could run as high as 29% with the near-unit root problem of forward discount. If the trend is included, rejection rates exceed 60%.

Nevertheless, the estimation of the $AR(p)$ model using the classical estimator of ordinary least square (OLS) is biased and quite large when the root is close to one (i.e. Mariott and Pope, 1954; Pantula and Fuller, 1985; Ledolter, 2009; Shaman and Stine, 1988). At the boundary of one (the unit root process), the OLS estimator converges to the true value at a much faster rate than other points of the parameter space, a non-uniform convergence of parameters of interest to limiting distribution. Furthermore, the unit-root distribution is heavily skewed toward the left. This behaviour carries over to the distributions of the regression t statistics. In studies conducted by Chan and Wei (1987, 1988), Chan (1988) and Phillips (1987), they showed that the distribution of

OLS estimator for root that is close to one is similar to limiting distribution at unit-root distribution.

The focus of this chapter is to understand the degree of persistency in forward discount in autoregressive (*AR*) process. As such, the focus of this chapter is two folds; to measure the persistency of forward discount based on confidence interval approach due to bias of OLS on point estimate of the root near unity and to determine whether there are changes in the persistency of the series between $I(0)$ and $I(1)$ process in the sample period.

Constructing the confidence interval for the sum of the *AR* coefficient is rather difficult (Basawa et al., 1991). Basawa et al. (1991) showed that the bootstrapped confidence interval is not valid as a measure of persistence for a nearly integrated process. This is also the case for the forward discount, where previous studies have provided evidence of near or borderline nonstationary (i.e., Maynard, 2006; Maynard and Phillips, 2001; Crowder, 1995; Baillie and Bollerslev, 1994a). To circumvent this problem, valid procedures of (Hansen, 1999; Romano and Wolf, 2001; Stock, 1991) were used in this chapter. In the context of changes in persistence analysis, we apply time varying (*TR*) autoregressive $AR(p)$ model proposed by Leybourne et al. (2007). This model allowed us to determine whether the forward discount undergoes regime shifts between $I(0)$ and $I(1)$ behavior with consistent change dates estimation.

6.3 Methodologies

Consider an $AR(p)$ process of variable y_t :

$$y_t = \mu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + e_t \quad (6.1)$$

for $t = 1, 2, \dots, T$, where the scalar measure of persistence is the sum of *AR* coefficients, $\alpha = \sum_{i=1}^p \alpha_i$. This measure is related to the cumulative impulse response (*CIR*)

$= 1/(1 - \alpha)$, where the scalar measure α is more informative than the largest root of the $AR(p)$ model in determining the persistence of the series (Andrews and Chen, 1994).

Parameter α can be estimated through ordinary least squares (OLS) by using an augmented Dickey-Fuller (ADF) (1979) regression model:

$$y_t = \mu' + \alpha y_{t-1} + \sum_{j=1}^k \beta_j \Delta y_{t-j} + e_t \quad (6.2)$$

where $\Delta y_t = y_t - y_{t-1}$ and $k = p - 1$. However, a problem arises in constructing the confidence intervals of α based on OLS estimation because the asymptotic distribution of the OLS estimator and its rate of convergence differ for the stationary and unit root cases. For $\alpha < 1$, the confidence interval can be constructed using the conventional asymptotic method; however, this method performs poorly in finite samples when α is near unity. As a result, the conventional bootstrap method is used in such cases. When α is near unity, however, the conventional bootstrap also fails to generate a confidence interval with the correct first-order asymptotic coverage (Basawa et al., 1991).

Several methods have been suggested to construct the above confidence interval of α . These methods include Hansen's (1999) grid bootstrap, Romano and Wolf's (2001) equitailed and symmetric subsampling and Stock's (1991) inversion of the ADF t -statistic. A recent study conducted by Mikusheva (2007) discussed the uniform validity of the above methods in constructing this confidence interval. She argued that the grid bootstrap and inverse ADF t -statistic provide uniform approximation and construct an asymptotically valid confidence set. However, the equitailed subsampling of Romano and Wolf (2001) is not asymptotically uniform. For symmetric subsampling of Romano and Wolf (2001), it generates asymptotically uniform coverage but rather

conservative, that is, the limit of the maximal coverage is higher than the declared level (Andrews and Guggenberger, 2009).

Another issue in constructing the confidence interval above arises when there is conditional heteroskedasticity that result in an incorrect asymptotic size.⁵⁰ A study conducted by Andrews and Guggenberger (2011) showed that with various confidence interval methods, only symmetric subsampling of Romano and Wolf (2001) has correct asymptotic size in the presence of conditional heteroskedasticity. For Hansen's (1999) grid bootstrap method and Stock's (1991) inversion of the ADF t -statistic, the confidence interval has correct asymptotic size only for conditional homoskedasticity.

Studies conducted by Andrews and Guggenberger (2009, 2010)) focusing on the issue of constructing confidence intervals in autoregressive model with root that are close to unity which they called as 'non-regular' models. They show that equitailed subsampling have substantial asymptotic distortions while symmetric subsampling have the correct asymptotic size when the root of the autoregressive root is near unity. The above studies highlighted the strengths and weaknesses of grid bootstrap method of Hansen (1999), symmetric subsampling of Romano and Wolf (2001) and inversion of the ADF t -statistic of Stock's (1991) to determine the confidence interval of α in the above equation. Given that no single method is superior, this chapter will apply all the methods above in determining the confidence interval of α .

Hansen (1999) argued that the grid bootstrap is asymptotically correct even in the context of a local-to-unity framework. Its method asymptotically controls Type I errors. Consider a grid values for α of $\alpha_i (i = 1, \dots, B)$ that covering $\hat{\alpha}$. The data-generating process for each α_i is estimated using Eq. (6.2) through restricted OLS where α restricted for each α_i . This restricted OLS parameter estimates with resampled

⁵⁰ See Andrews and Guggenberger (2011).

restricted OLS residuals used to build up a large number of pseudo samples of $B = 1999$ for each α_i .⁵¹ With each α_i of the pseudo samples giving us the t_{α_i} -statistic, the t_{α_i} -statistic was then sorted out, which gave us an empirical distribution. When then calculated the 0.025 and 0.975 quantiles of the t_{α_i} -statistic for each α_i from this empirical distribution. This gave us the confidence interval for α , where the upper bound of 95% is the α_i grid value of $(\hat{\alpha} - \alpha_i)/s(\hat{\alpha}) = t_{\alpha_i,0.025}$ and the lower bound is the grid value of α_i with $(\hat{\alpha} - \alpha_i)/s(\hat{\alpha}) = t_{\alpha_i,0.975}$.

Another approach to constructing the confidence interval is the subsampling procedure (Romano and Wolf, 2001). Romano and Wolf (2001) offered two procedures, two-sided equal-tailed and two-sided symmetric confidence intervals. However, as mentioned earlier, recent studies have shown that the equal-tailed has asymptotic size distortion while the symmetric confidence interval performs well even in the existence of conditional heteroskedasticity. The procedure is used to estimate by using OLS in smaller blocks depending on the sample size. Let b_t for $t = 1, \dots, T - b + 1$ be the block of size b ; we computed the t -statistics for α of this subsample of $\tau_b(\hat{\alpha}_{b,t} - \hat{\alpha})/\hat{\sigma}_{b,t}$ where τ_b is an appropriate normalizing constant and $\hat{\sigma}_{b,t}$ is OLS estimate of α for the t th block of size b , $\hat{\sigma}_{b,t} = b^{1/2}s(\hat{\sigma}_{b,t})$ and $\tau_b = b^{1/2}$. For the two-sided symmetric subsampling interval, the empirical approximating distribution for the subsample t -statistics is:

$$L_{b,|\cdot|}(x) = \frac{1}{T - b + 1} \sum_{t=1}^{T-b+1} 1\{\tau_b|\hat{\alpha}_{b,t} - \hat{\alpha}|/\hat{\sigma}_{b,t} \leq x\} \quad (6.3)$$

⁵¹ As suggested by Hansen (1999).

and the 95% two-sided symmetric confidence interval is:

$$[\hat{\alpha} - (1/\tau_T)s(\hat{\alpha})c_{b,| \cdot |,0.05}, \hat{\alpha} + (1/\tau_T)s(\hat{\alpha})c_{b,| \cdot |,0.05}] \quad (6.4)$$

where $c_{b,| \cdot |,0.05}$ is the 0.05 quantiles of the empirical distribution.

For the subsampling method to be consistent, the block size b needs to tend toward infinity with the sample size T (Romano and Wolf, 2001). This is because when b is close to sample size T , the subsampling distribution will result in under-coverage of the subsampling confidence intervals. Another issue arises when b is too small, so the intervals will under- or over-cover the subsampling confidence intervals. To overcome these issues, we implemented Romano and Wolf's (2001) algorithm to minimize confidence interval volatility.

In Stock (1991), the confidence interval of α is constructed from the hypothesis test that $\alpha = 1$ form augmented Dickey-Fuller (ADF) test of Eq. (6.2). However, the ADF statistic's distribution is highly non-normal result in the construction of asymptotic confidence intervals as a point estimate \pm two standard errors is not suitable. Additionally, one cannot rely on traditional asymptotic theory since it is discontinue at $\alpha = 1$. Alternatively, we can use local-to-unity asymptotic theory that models the true value of α being in a $1/T$ neighborhood of one:

$$\alpha = 1 + c/T \quad (6.5)$$

where c is a fixed constant. Under local-to-unity framework above, the limiting distribution of the ADF statistics depends on c . The confidence interval of the ADF statistics that are constructed is then inverted to yield confidence interval of c , or equivalently, given T , confidence interval of α . In particular, a $100(1 - 0.05)\%$ confidence interval based on realized value of the ADF statistic, say $\hat{t}(c)$ is given by:

$$\{c: f_{0.025}(c) \leq \hat{\tau}(c) \leq f_{0.975}(c)\} \quad (6.6)$$

where $f_{0.025}(c)$ and $f_{0.975}(c)$ are the percentiles of τ as a function of c . As f is strictly monotone in c , we can invert the preceding expression to give:

$$\{c: f_{0.025}^{-1}(c) \leq \hat{\tau}(c) \leq f_{0.975}^{-1}(c)\} \quad (6.7)$$

the 95% confidence interval of the ADF t -statistic inversion of Stock (1991).

The analysis of change in persistence is based on Leybourne et al.'s (2007) time varying $AR(p)$ model, which allows us to determine multiple changes of persistence of the forward discount. In the time series literatures, Leybourne et al.'s model is the only model that is valid in the presence of multiple changes in persistence. As shown in Leybourne et al.'s (2007), a single change in persistence of Kim (2000), Harvey et al. (2006) and Leybourne et al. (2006) are inconsistent against process which display multiple change in persistence. The model suggested by Leybourne et al.'s (2007) is:

$$y_t = d_t + u_t \quad (6.8)$$

$$u_t = \rho_i u_{t-1} + \sum_{j=1}^{k_i} \phi_{i,j} \Delta u_{t-j} + \varepsilon_t \quad t = 1, \dots, T \quad (6.9)$$

where $k_i = p_i - 1$ and deterministic kernel $d_t = z_t' \beta$ is assumed to be constant or $z_t = 1$ and $\beta = \beta_0$ in this study.⁵² The u_t in (6.8) and (6.9) is taken to be time varying $AR(p)$ process and $i = 1, \dots, m + 1$ where m represents the number of changes in persistence. The model in (6.8) and (6.9) permits that the dominant AR root, ρ_i , and the lag coefficients, $\phi_{i,j}$, differ across the $m + 1$ separate regimes. The null hypothesis is that y_t is an $I(1)$ process throughout the sample period or $\rho_i = 1$ with the alternative that y_t is subject to one or more regime shifts between $I(0)$ and $I(1)$ behaviour. Thus,

⁵² Since the forward discount is a non-trending series.

under alternatives, ρ_i is subject to $m \geq 1$ unknown persistence change points that give rise to $m + 1$ segments with change point fraction given by $\tau_1 < \tau_2 < \dots < \tau_{m-1} < \tau_m$.

Leybourne et al. (2007) define the fraction $\tau \in (\lambda, 1)$, for a given $\lambda \in (0, 1)$ and the test is based on the local *GLS* de-trended *ADF* unit root statistics. The test uses the sample observation between λT and τT , called $DF_G(\lambda, \tau)$ through standard *t*-statistics associated with $\hat{\rho}_i$ based on the fitted regression of:

$$\Delta y_t^d = \hat{\rho}_i y_{t-1}^d + \sum_{j=1}^{k_i} \hat{b}_{i,j} \Delta y_{t-j}^d + \hat{\varepsilon}_t \quad t = \lambda T, \lambda T + 1, \dots, \tau T \quad (6.10)$$

where $y_t^d \equiv y_t - z_t' \hat{\beta}$, with $\hat{\beta}$ is obtained through OLS by regressing $y_{\lambda, T}$ on $z_{\lambda, T}$, where $y_{\lambda, T} \equiv (y_{\lambda, T}, y_{\lambda, T+1} - \bar{\alpha} y_{\lambda, T}, \dots, y_{\tau T} - \bar{\alpha} y_{\tau T-1})'$, $z_{\lambda, T} \equiv (z_{\lambda, T}, z_{\lambda, T+1} - \bar{\alpha} z_{\lambda, T}, \dots, z_{\tau T} - \bar{\alpha} z_{\tau T-1})'$ with $\bar{\alpha} = 1 + \bar{c}/T$ and $\bar{c} = -10$. In this study, τ is set to 0.20 following Leybourne et al. (2007). The test suggested by Leybourne et al. (2007) is based on doubly recursive sequence of *DF* type of unit root statistics:

$$M = \inf_{\lambda \in (0, 1)} \inf_{\tau \in (\lambda, 1)} DF_G(\lambda, \tau) \quad (6.11)$$

with corresponding estimator $(\hat{\lambda}, \hat{\tau}) \equiv \arg \inf_{\lambda \in (0, 1)} \inf_{\tau \in (\lambda, 1)} DF_G(\lambda, \tau)$. The M test yields the start and end points of the interval $[\hat{\lambda}, \hat{\tau}]$ of the first $I(0)$ regime over the whole sample. Further $I(0)$ regimes can be detected sequentially by applying the M statistic to each of the resulting subintervals $[0, \hat{\lambda}]$ and $[\hat{\tau}, 1]$. Continuing in this way, all $I(0)$ regimes can be determined with their start and end points.

6.4 Findings⁵³

In second and third columns of Table 6.1, the OLS estimate of α of Eq. (6.2) with the lag k determined by modified Akaike information criteria (MAIC) proposed by Ng and Perron (2001) are reported. This method is used to determine the lag taking into account the fact that the bias in the sum of autoregressive coefficient is highly dependent on k . The OLS point estimate of α shows greater than 0.90, a highly persistence process for German mark, Italian lira and U.S. dollar. This result is expected given the findings of previous studies. For example, in Sakoulis et al. (2010), they found $\alpha > 0.90$ for German mark, Japanese yen and U.K. pound. For Japanese yen, our result is quite similar to Hai et al. (1997) where $\alpha = 0.923$.

However, these point estimates of OLS are biased and not really useful as mentioned earlier. In order to provide alternative measures of persistence, we calculate Hansen's (1999) grid bootstrap, Romano and Wolf's (2001) symmetric subsampling and Stock's (1991) inversion of the ADF t -statistic for each G7 currencies' forward discount.

In the fourth column of Table 6.1, we have a 95% confidence interval of the grid bootstrap method (Hansen, 1999). All currencies have a lower bound of less than 0.90. However, the upper bound of the 95% confidence interval includes unity for German mark and Italian lira that is consistent with unit root process.

Stock's (1991) inverse ADF t -statistic 95% confidence interval is reported in the fifth column of Table 6.1. For the lower bound of 95% confidence interval, the results are quite similar to grid bootstrap except for Canadian dollar, French franc and Japanese

⁵³ The grid bootstrap 95% confidence interval is conducted using GAUSS code written by Bruce Hansen and available at <http://www.ssc.wisc.edu/~bhansen/>
The inverse ADF t -statistics 95% confidence interval is conducted using GAUSS code written by James Stock and available at <http://scholar.harvard.edu/stock/publications/confidence-intervals-largest-autoregressive-root-macroeconomic-time-series>
The symmetric subsampling 95% confidence interval is conducted using GAUSS code written by David Rapach and available at <http://sites.slu.edu/rapachde/home/research>

yen where large differences in estimation comparing to grid bootstrap result. Two of the currencies lower bound, German mark and Italian lira, is equal or greater than 0.90 that indicate these currencies' forward discount display a high degree of persistence. Based on the upper bound of 95% confidence interval, the results are similar to grid bootstrap where German mark and Italian lira are consistent with unit root process.

Table 6.1: Point estimate and 95% confidence intervals for measure of forward discount persistency

	k_{MAIC}^a	$\hat{\alpha}_{OLS}^b$	Grid-bootstrap 95% CI	Inverse ADF 95% CI	Symmetric subsampling 95% CI
Canadian dollar	10	0.859	[0.801, 0.941]	[0.713, 0.959]	[0.801, 0.917]
French franc	9	0.849	[0.779, 0.970]	[0.872, 0.983]	[0.744, 0.955]
German mark	10	0.931	[0.893, 1.007]	[0.900, 1.004]	[0.867, 0.995]
Italian lira	10	0.927	[0.873, 1.034]	[0.947, 1.013]	[0.826, 1.028]
Japanese yen	12	0.877	[0.823, 0.957]	[0.686, 0.969]	[0.818, 0.935]
U.S. dollar	12	0.907	[0.863, 0.977]	[0.895, 0.982]	[0.854, 0.960]

^aLag length k is based on (Ng and Perron, 2001) of modified Akaike information criteria (MAIC)

^bOLS estimate for the sum of the AR coefficients in Eq. (6.2)

In the last column of Table 6.1, we report the Romano and Wolf's (2001) symmetric sub-sampling 95% confidence interval for α . For the most part, the confidence interval of α is quite similar to the grid-bootstrap confidence interval. The lower bounds are equal to or greater than 0.74 for all currencies. For the upper bound,

Italian lira included unity, another indication that this currency is consistent with unit root process.

All procedures of the 95% confidence interval suggest large degrees of uncertainty regarding the degree of persistence of the forward discount. The result also shows that the forward discount is quite persistent with some upper bounds of the 95% confidence interval, including unity of which is consistent with the unit root process. This uncertainty regarding the degree of persistence in the explanatory variable of Eq. (2.7) suggests possible size distortion in standard tests of $\beta_2 = 1$. As shown in Cavanagh et al. (1995), regression such as Eq. (2.7) exhibits large size distortions if the regressor has local-to-unit roots.

In Table 6.2, the result of result of M -test for multiple changes in persistence of Leybourne et al. (2007) are reported. In the third column of Table 6.2, we report the M statistics, while the last two columns are the respective beginning and end of $I(0)$ regimes identified by the procedure. The M -test is initially applied over the whole sample, detecting an interior $I(0)$ regime. If the null of $I(1)$ is rejected based on the whole sample, the presence of subsequent $I(0)$ regime will be tested against each of the resulting subsamples. Continuing in this way, all $I(0)$ regimes can be determined by their start and end points. No further testing is applied if the test fails to detect any interior $I(0)$ regime.

In order to understand the procedure, we use French franc as an example. The test is initially applied over the whole sample (1976:01 – 1998:12) of French franc, detecting an interior $I(0)$ regime between 1993:04 and 1998:07, for which the unit root null is rejected at the 1% level. This represents the most prominent $I(0)$ regimes in the data. The test is then applied over 1976:01-1993:03 and the test rejects the null at the 10% level, identifying the second $I(0)$ regime between 1979:02 and 1983:02. This

represents, in turn, the most prominent $I(0)$ region within this subsample. The search for further stationary regime continues by applying the test over the sample 1983:03-1993:03, which yields a third $I(0)$ regime corresponding to the period 1990:04-1992:10. Hence, the procedure detects a total of 3 $I(0)$ regimes for the whole sample of French franc. For Japanese yen, the test applied over the whole sample (1978:06-2005:12) fails to detect any $I(0)$ regime. As such, no further test is conducted for Japanese yen.

Overall, most of the currencies involved experienced changes in persistence between $I(0)$ and $I(1)$ regimes throughout the sample period, with exception to Japanese yen. This finding is interesting given by the fact that in previous studies, it was concluded that forward discount is a borderline stationary process (e.g. Hall et al. (1997) argue that U.K. pound, French franc and Japanese yen is borderline stationarity (α close to unity, if $\alpha = 1$ it is nonstationary) with α is 0.812, 0.798 and 0.923 respectively). Furthermore, the 95% confidence interval result shows uncertainty regarding the degree of persistence of the forward discount. In the context of unit root testing, the augmented Dickey-Fuller (ADF) test that is commonly applied in testing for unit root in forward discount will not provide consistent testing procedure since it cannot discriminate between fixed $I(1)$ process and persistence change series. This is due to the fact $I(1)$ part of the series will dominate asymptotically (see Chong, 2001; Taylor, 2005). Figure 6.1 gives a graphical representation of the result, where the shaded region represents $I(0)$ regimes.

Table 6.2: Multiple persistent change test results

	Period	M	I(0) Periods	
			Start	End
Canadian dollar	1976:01 - 2005:12 (full sample)	-5.174***	1979:10	2005:12
	1976:01-1979:09 ^a	-5.260*	1978:07	1979:05
	1976:01-1978:06 ^a	-5.872*	1977:11	1978:05
French franc	1976:01 – 1998:12 (full sample)	-5.733***	1993:04	1998:07
	1976:01-1993:03	-4.723*	1979:02	1983:02
	1983:03-1993:03	-5.850***	1990:04	1992:10
German mark	1976:01 - 1998:12 (full sample)	-6.783***	1976:02	1980:12
	1981:01-1998:12	-4.776**	1984:07	1988:07
Italian lira	1976:01 - 1998:12 (full sample)	-5.508**	1981:05	1990:10
	1976:01-1981:04	-7.428***	1976:04	1977:06
	1977:07-1981:04	-9.331***	1978:03	1979:03
Japanese yen	1978:06 - 2005:12 (full sample)	-3.888	-	-
U.S. dollar	1976:01 – 2005:12 (full sample)	-6.294***	1980:03	1992:08
	1976:01-1980:02	-25.72***	1978:09	1979:07

***, ** and * denotes significance at the 1%, 5% and 10% respectively

This finding complements Sakoulis et al.'s (2010) study, which argued that the high persistence of the forward discount is exaggerated by the presence of a break in the mean of the series. In Sakoulis et al. (2010), with structural breaks, they found a significant drop in $AR(1)$ process of forward discount. The estimate of α based $AR(1)$ model for Canadian dollar, French franc, German mark and Italian lira are 0.711, 0.415,

0.666 and 0.558 respectively. Thus, this chapter's finding of multiple changes in persistence adds to the current literature on the degree of persistency of the forward discount.

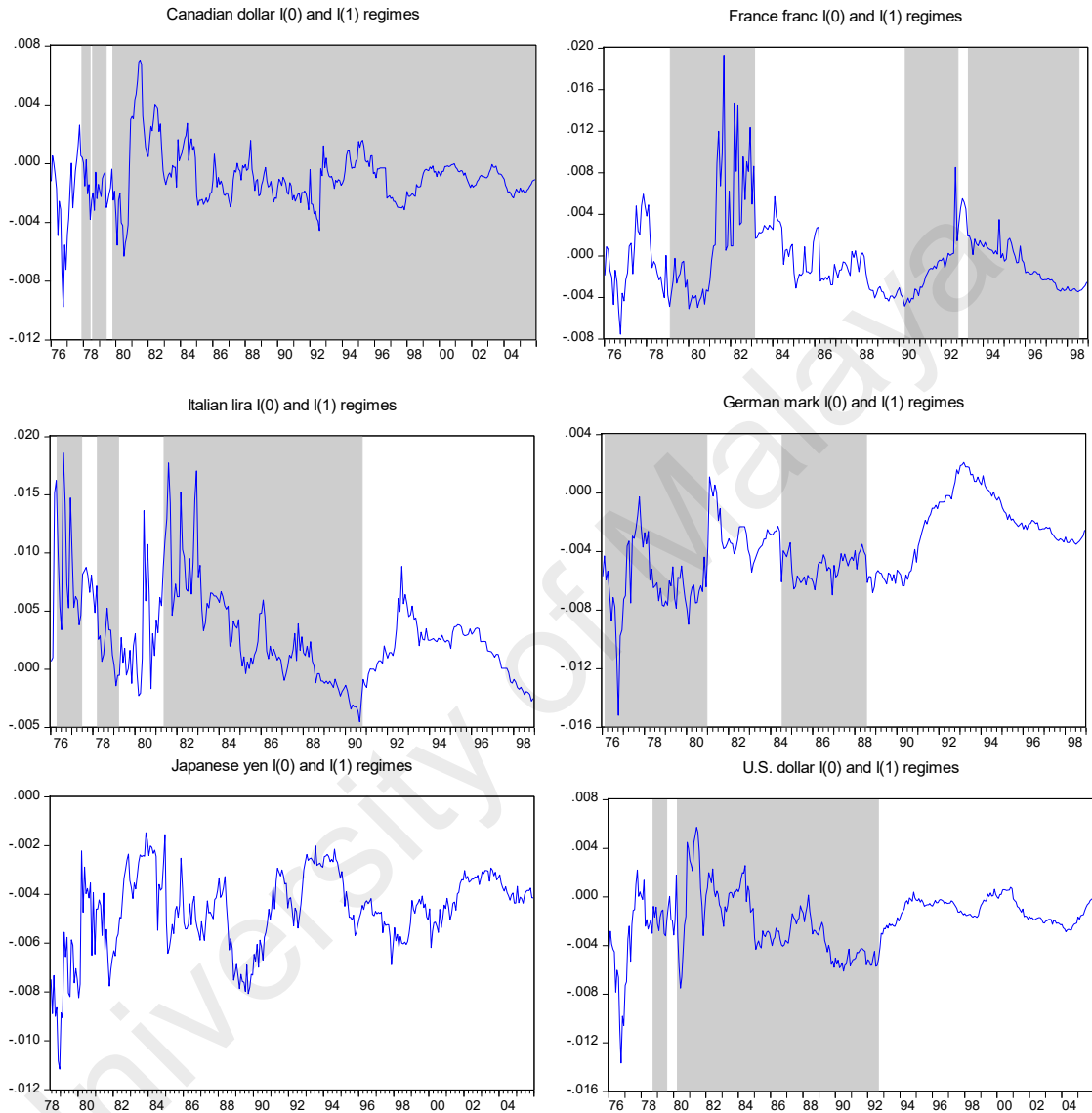


Figure 6.1: $I(0)$ and $I(1)$ regimes of forward discount

6.5 Conclusion

In this chapter, the study investigated the persistency of the forward discount based on the confidence interval approach and multiple changes in persistence based on time varying autoregressive model. Given that the conventional confidence interval approach

fails to estimate when the series is a near unit root process, which is the case for the forward discount, this chapter used Hansen's (1999) grid bootstrap, Romano and Wolf's (2001) symmetric subsampling procedures and Stock's (1991) inversion of the ADF t -statistic. These methods provide uniform approximation and construct asymptotic confidence set, with symmetric subsampling also provide correct asymptotic size in the present of conditional heteroskedasticity.

Based on the G7 countries' currencies used, our results regarding the confidence interval show a large degree of uncertainty in persistence where the upper bound result includes unity in some currencies, which is consistent with the unit root process. The results of the multiple changes in persistence test highlight that the forward discount undergoes change in persistence between stationary and nonstationary regime for most of the currencies.

Overall, our findings highlight that the forward discount undergoes changes in persistent between $I(0)$ and $I(1)$ process throughout the entire sample period. This has important implications in the context of FRUH testing. This finding shows that forward discount indeed has nonstationary components as proposed by Evan and Lewis (1995) and Crowder (1994), rather than borderline nonstationarity process of forward discount (i.e. Liu and Maynard, 2005; Maynard, 2006). Thus, future research in the context of potential bias in the empirical inference procedure can provide further understanding of the role of changes in persistence of forward discount in testing for FRUH.

Chapter 7: Conclusion, Implications and Limitations

7.1 Introduction

The aim of this study is to examine the time series properties of the forward discount, in particular the persistency of the series. Understanding the persistence of the forward discount has become the centre of discussion recently due to issues raised on the econometric inference of FRUH testing that leads towards the forward bias puzzle. In this econometric issue, the central idea is that highly persistent behaviour of the forward discount results in bias or size distortion in FRUH testing. In modelling the persistence process of the forward discount, two models are used; long memory and the root near unity of the autoregressive (*AR*) process.

Historically, the exact stochastic process of forward discount is undecided. Some studies favour a stationary process, while others favour a nonstationary process. The conflicting findings have resulted in recent studies that employ an intermediate stochastic process between stationary and nonstationary, a persistence process of long memory or root near unity.

Motivated by the recent studies of Choi and Zivot (2007) and Sakoulis et al. (2010) that focus on the persistency of the forward discount, this study extends the discussion by employing econometric methods that are relevant in understanding the persistency of the series. For that purpose, we use G7 countries' currency since they are extensively used in previous studies of FRUH (see Sarno, 2003; Engle, 1996).

7.2 Summary of The Thesis

This thesis is divided into seven chapters; the objectives of the study are addressed in Chapters 4, 5 and 6. In Chapter 4, the focus is on the previous findings regarding long memory in the forward discount. Given that structural breaks or regime switching may induce long memory behaviour, previous findings related to long memory are in doubt, as structural breaks are suspected in the forward discount as monetary authorities intervene in the foreign exchange market to counter business cycles. Overall, the results from Chapter 4 favour a true long memory process in the G7 currencies used in this study. Even though structural breaks fail to explain the long memory behaviour found in the forward discount, adjusting for breaks did reduce the persistence of long memory parameter estimates. Thus, we might suspect that long memory findings in previous studies involve a true long memory process, but the estimations are overestimated due to structural breaks.

In Chapter 5, the focus is on the issue of mixed finding of stationarity in the forward discount. Even though recent studies are no longer focusing on the extreme property of either stationary or nonstationary, this chapter offers new arguments on the previous mixed findings. This issue is important given that numerous studies have previously reported nonstationarity in the forward discount, which does not seem particularly plausible in an economic sense.

In previous studies, testing for stationarity relies on Dickey-Fuller-type testing, or any classical unit root test assuming that the data are linear. This chapter provides theoretical arguments that suggest the forward discount is a nonlinear process. The nonlinearity in the forward discount might be due to inter alia transaction cost, which creates a 'band of inaction' where inside the band no adjustment in deviation from equilibrium takes place. However, the process becomes mean reverting outside the band

due to benefits of arbitrage exceeding the cost. In Chapter 5, we rely on the symmetric threshold autoregressive model with an autoregressive unit root developed by Caner and Hansen (2001). The model allows for an inner no-arbitrage band with small disequilibria while capturing mean reversion to shocks outside the no-arbitrage band consistent with the transaction cost hypothesis argument of nonlinearities.

All the currencies show strong evidence of nonlinearities due to threshold. Thus, all the currencies involved can be modelled as two-regime symmetric TAR. Further analysis provides clear evidence that five currencies possess one nonstationary regime and one stationary regime. This is in line with the argument of transaction cost where inside the band no adjustment in deviation from equilibrium takes place when the deviation follows a unit root process. However, outside the band, the benefit of arbitrage exceeds the cost, where the process switches abruptly to become mean reverting towards the transaction cost band. Overall, the forward discount is shown to be globally stationary albeit with high persistency since most of the observations lie inside the dominant unit root regime. The inside band (unit root regime) is dominant in all currencies; it fluctuates between 61% and 85% in the observations. This might explain the findings of nonstationarity in previous studies of the forward discount. We conclude Chapter 5 with the forecasting performance of the linear $AR(k)$ model against the threshold model used in this chapter. Based on root mean squared error (RMSE) statistic, the TAR model outperforms the linear $AR(k)$ model in 1, 3, 6, 9 and 12 month forecasts in most currencies.

In Chapter 6, the focus is on modelling the persistence of the forward discount based on the autoregressive $AR(p)$ model with the root near unity. In this context, the forward discount is modelled as a 'borderline nonstationarity' process (Maynard, 2006). This chapter revisits the $AR(p)$ modelling of the forward discount based on (1) the point estimate of the $AR(p)$ model using OLS being biased and quite large when the root is

close to one and (2) whether the forward discount indeed has nonstationary components, as Evan and Lewis (1995) and Crowder (1994) propose, or is a borderline nonstationarity process of forward discount (i.e. Liu and Maynard, 2005; Maynard, 2006).

The approaches in Chapter 6 are twofold; using the confidence interval approach rather than a point estimate of the largest root in the $AR(p)$ model and using multiple change persistence tests suggested by Leybourne et al. (2007). The result of the confidence interval shows a large degree of uncertainty in persistence where the upper bounds results includes unity in some currencies, which is consistent with the unit root process. Interestingly, the result of multiple changes in persistence tests shows that the forward discount undergoes change in persistence between stationary and nonstationary regimes for most of the currencies involved. This finding shows that the forward discount indeed has nonstationary components, as Evan and Lewis (1995) and Crowder (1994) propose, rather than a borderline nonstationarity process of forward discount (i.e. Liu and Maynard, 2005; Maynard, 2006).

7.3 Implications

Three major findings of this study apply the econometric issue in explaining the forward bias puzzle. First, this study compiles strong evidence that the forward discount is a true long memory process, where structural breaks fail to explain long memory behaviour in the forward discount. Second, the forward discount undergoes multiple changes in persistence between the $I(0)$ and $I(1)$ processes and, last, the forward discount might be a nonlinear process based on the argument regarding transaction cost.

In testing for FRUH, most studies have relied on regression Eq. (2.7). Since the exchange rate return or the dependent variable is stationary, in the context of the econometric inference the problem lies in the persistent behaviour of the forward discount. It has been understood since Mankiw and Shapiro (1986) that persistence in the regressor may lead to size distortion in Eq. (2.7). In the context of the econometric literature, most studies focus on the case of near unit root regressors as a persistent process. If the largest root of the regressor is close to but not equal to unity, the usual practice is to adjust the critical value to preserve the correct test size. Thus, regression of Eq. (2.7) is still compatible with stationarity in the exchange return of the dependent variable. However, in this study we find that the forward discount actually possesses a unit root component where it undergoes multiple persistence changes in between the $I(0)$ and $I(1)$ processes. This suggests modelling the forward discount as the largest root that close to unity is inappropriate.

If the forward discount is truly a unit root process, Eq. (2.7) will result in an unbalanced regression where the return is an $I(0)$ process and the forward discount is an $I(1)$ process (see Liu and Maynard, 2005). Furthermore, when the regressor has unit root, the true regression coefficient is forcibly equal to zero (see Maynard and Shimotsu, 2009). Torous et al. (2004) show that inference in regressions like Eq. (2.7) depend critically on the assumed stochastic properties of regressors whether the variable is assumed to be nonstationary or stationary. Thus, for future studies, FRUH testing requires taking into account that the forward discount undergoes multiple changes in persistence between the $I(0)$ and $I(1)$ processes rather than modelling as the root near unity process. However, it is not clear whether current econometric techniques can handle such a scenario.

Note that long memory is another manifestation of the persistence process despite root near unity. However, less attention has been paid when the regressor displays long memory behaviour. As Maynard and Shimotsu (2009) note, the remedies employed in the context of near unit roots do not necessarily carry over to the long memory case. With the exchange rate return is stationary in Eq. (2.7), an unbalanced regression becomes unavoidable when the forward discount has nonstationary long memory or $d > 0.5$ (i.e. Maynard and Shimotsu, 2009). As reported in Chapter 4, most of the currencies' long memory estimate based on MLP and FELW before adjustment of structural breaks show evidence of nonstationary long memory. Even though after adjustment for breaks the persistence in long memory declines, we still find evidence of nonstationary long memory in some of the currencies. For future research, filtering the forward discount based on the order of integration (long memory) might reduce the issue of unbalanced regression. However, further analysis is needed since structural breaks also exist in the series where filtering based on the order of integration ignores the existence of breaks.

In the context of nonlinearities in the forward discount, our finding suggests the possibility that Fama's (1984) regression-based test is clearly miss-specified due to the assumption of linearity in the regressor. Thus, future research is required to model FRUH testing in a nonlinear form.

7.4 Limitations

This study has a few limitations. In the context of spurious long memory in Chapter 4, the sample used in this study is considered small in the long memory literature. In the long memory literature, the samples used normally involve thousands of observations.

Furthermore, the semiparametric methods used critically rely on the periodogram ordinates or bandwidth selection of m . This balances a trade-off between variance and bias. If m is too large or too small, the outcome of the estimation may wrongly suggest a certain degree of persistence. In normal practice, many empirical researchers typically opt for a grid of bandwidth values and then plot the estimates against different values of m (see Taqqu and Teverovsky, 1996). However, the plot only works if the sample involved is really large. With a small sample, we will experience fluctuation in d when plotted against m . The d will normally be stable if m is large enough, $m = T^\alpha$, or large in T . This can be avoided if daily data are used. However, to stay within the framework of the original FRUH testing, we opt for monthly data following Choi and Zivot (2007). To avoid such pitfall, in Chapter 4 we report the semiparametric results based on the difference level of m that is suggested in previous studies.

In the context of nonlinearity, we focus only on the existence of threshold as evidence of nonlinearity. In actual fact, nonlinearity can appear in many forms. In previous studies, nonlinearities in exchange rate have been modelled as smooth transition autoregressive (STAR) (i.e. Granger and Terasvirta, 1993) exponential STAR (ESTAR) (i.e. Panos et al., 1997) and logistic STAR (LSTAR) (i.e. Panos et al., 1997). However, these models assume stationarity prior to fitting nonlinearity. By assuming stationarity, the models do not reconcile the nonlinear adjustment with the evidence of unit root in the forward discount that previous studies have reported. Given that the only model available in the econometric literature that tackles both nonlinearity and/or nonstationary is Caner and Hansen's (2001) model, we have to assume that the regime change is discrete rather than smooth even though the smooth regime change has gained popularity (see Terasvirta, 1994).

It is noted earlier that none of the confidence interval approaches used in Chapter 6 is superior. Every model has some weaknesses and strengths. Even though we

find multiple changes in persistence, the model cannot ascertain the degree of persistence in $I(0)$ regimes. The $I(0)$ regimes might also be highly persistent, but stationary (root close to unity).

University of Malaya

REFERENCES

- Adler, M., & Dumas, B. (1983). International portfolio choice and corporation finance: A synthesis. *Journal of Finance*, 38, 925-984.
- Aggarwal, R., Lucey, B. M., & Mohanty, S. K. (2009). The forward exchange rate bias puzzle is persistent: Evidence from stochastic and nonparametric cointegration tests. *The Financial Review*, 44, 625-645.
- Agiakloglou, C., Newbold, P., & Wohar, M. (1993). Bias in an estimator of the fractional difference parameter. *Journal of Time Series Analysis*, 14(3), 235-246.
- Alexander, C.O., & Johnson, A. (1992). Are foreign exchange market really efficient? *Economic Letters*, 40, 449-453
- Alvarez, F., Atkeson, A., & Kehoe, P. J. (2002). Money, interest rates, and exchange rates in endogenously segmented markets. *Journal of Political Economy*, 110(1), 73-112.
- Andrews, D. W., & Chen, H. Y. (1994). Approximately median-unbiased estimation of autoregressive models. *Journal of Business & Economic Statistics*, 12(2), 187-204.
- Andrews, D. W., & Guggenberger, P. (2009). Hybrid and size-corrected subsampling methods. *Econometrica*, 77(3), 721-762.
- Andrews, D. W., & Guggenberger, P. (2010). Applications of subsampling, hybrid and size-correction methods. *Journal of Econometrics*, 158, 285-305.
- Andrews, D. W., & Guggenberger, P. (2011). A conditional-heteroskedasticity-robust confidence interval for the autoregressive parameter. *Cowles Foundation Discussion Paper No. 1812*, 1-20.
- Bacchetta, P., & van Wincoop, E. (2010). Infrequent portfolio decisions: A solution to the forward discount puzzle. *American Economic Review*, 100(3), 870-904.
- Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1-22.
- Baillie, R. T. (1996). Long memory process and fractional integration in econometrics. *Journal of Econometrics*, 73, 5-59.
- Baillie, R. T. (2011). Possible solutions to the forward bias paradox. *Journal of International Financial Markets, Institutions and Money*, 21, 617-622.
- Baillie, R. T., & Bollerslev, T. (1989). Common stochastic trends in a system of exchange rates. *The Journal of Finance*, 44(1), 167-181.

- Baillie, R. T., & Bollerslev, T. (1994a). Cointegration, fractional cointegration and exchange rate dynamics. *Journal of Finance*, 49, 737-745.
- Baillie, R. T., & Bollerslev, T. (1994b). The long memory of the forward premium. *Journal of International Money and Finance*, 13(5), 555-571.
- Baillie, R. T., & Bollerslev, T. (2000). The forward premium anomaly is not as bad as you think. *Journal of International Money and Finance*, 471-488.
- Barkoulas, J., Baum, C. F., & Chakraborty, A. (2003). Forward premiums and market efficiency: Panel unit-root evidence from the term structure of forward premiums. *Journal of Macroeconomics*(25), 109-122.
- Barnhart, S. W., McKnown, R., & Wallace, M. S. (1999). Non-informative tests of the unbiased forward exchange rate. *Journal of Financial and Quantitative Analysis*, 34, 265-91.
- Basawa, I. V., Mallik, A. K., McCormick, W. P., Reeves, J. H., & Taylor, R. L. (1991). Bootstrapping unstable first-order autoregressive processes. *Annals of Statistics*, 19, 1098-1101.
- Basci, E., & Caner, M. (2005). Are real exchange rates nonlinear or nonstationary? Evidence from a new threshold unit root test. *Studies in Nonlinear Dynamics & Econometrics*, 9(4), 1-19.
- Baum, C. F., Barkoulas, J. T., & Caglayan, M. (2001). Nonlinear adjustment to purchasing power parity in the post-Bretton Woods era. *Journal of International Money and Finance*, 20(3), 373-399.
- Bec, F., Ben Salem, M., & MacDonald, R. (2002). Real exchange rates and real interest rate: A nonlinear perspective. *Boston College: CREST-ENSAE*.
- Bekaert, G. (1995). The time-variation of expected returns and volatility in foreign exchange markets. *Journal of Business and Economic Statistics*, 13(4), 397-408.
- Bekaert, G., & Hodrick, R. J. (1992). Characterizing predictable components in excess returns on equity and foreign exchange markets. *Journal of Finance*, 47, 467-509.
- Bekaert, G., & Hodrick, R. J. (2001). Expectations hypothesis tests. *Journal of Finance*, 56(38), 1357-1394
- Bilson, J. F. (1981). The 'speculative efficiency' hypothesis. *The Journal of Business*, 54(3), 435-451.
- Bilson, J. F., & Levich, R. M. (1977). A test of the efficiency of the forward exchange market. *New York University Working Paper No. 77-61*.
- Borensztein, E. R. (1987). Alternative hypotheses about the excess return on Dollar assets. *International Monetary Fund Staff Papers*, 34, 29-59.

- Caner, M., & Hansen, B. E. (2001). Threshold autoregression with a unit root. *Econometrica*, 69(6), 1555-1596.
- Cavanagh, C.L., Elliott, G., & Stock, J.H. (1995). Inference in models with nearly integrated regressors. *Econometric Theory*, 11, 1131-1147
- Chakraborty, A., & Haynes, S. E. (2008). Econometrics of the forward premium puzzle. *Economics Bulletin*, 6(4), 1-17.
- Chan, N. H. (1988). The parameter inference for nearly nonstationary time series. *Journal of the American Statistical Association*, 83(403), 857-862
- Chan, N. H., & Wei, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *The Annals of Statistics*, 15(3), 1050-1063
- Chang, S. S. (2011). On the (In)feasibility of covered interest parity as a solution to the forward bias puzzle. *Journal of International Financial Markets, Institutions and Money*, 21, 611-616.
- Chinn, M. D. (2006). The (partial) rehabilitation of interest rate parity in the floating rate era: Longer horizons, alternative expectations, and emerging markets. *Journal of International Money and Finance*, 25, 7-21.
- Chinn, M. D., & Meredith, G. (2004). Monetary policy and long-horizon uncovered interest parity. *IMF Staff Papers*, 51(3), 409-430.
- Chinn, M. D., & Quayyum, S. (2012). Long horizon uncovered interest parity re-assessed. *NBER Working Paper No. w18482*.
- Choi, K., & Zivot, E. (2007). Long memory and structural changes in the forward discount: An empirical investigation. *Journal of International Money and Finance*, 342-363.
- Chong, T. T-L. (2001). Structural changes in AR(1) models. *Econometric Theory*, 17, 87-155.
- Clarida, R. H., & Taylor, M. P. (1997). The term structure of forward exchange premiums and the forecastability of spot exchange rates: correcting the errors. *The Review of Economics and Statistics*, LXXIX(3), 353-361.
- Clarida, R. H., Sarno, L., Taylor, M. P., & Valente, G. (2003). The out-of-sample success of term structure models as exchange rate predictors: A step beyond. *Journal of International Economics*, 60, 61-83.
- Crowder, W. J. (1992). Spot and forward exchange rates and the efficiency of foreign exchange markets. University of Texas at Arlington.
- Crowder, W. J. (1994). Foreign exchange market efficiency and common stochastic trends. *Journal of International Money and Finance*, 13(5), 551-564.

- Crowder, W. J. (1995). Covered interest parity and international capital market efficiency. *International Review of Economics and Finance*, 4(2), 115-132.
- Deegan Jr, J. (1976). The consequences of model misspecification in regression analysis. *Multivariate Behavioral Research*, 11(2), 237-248
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427-431.
- Diebold, F. X., & Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics*, 131-159.
- Dominguez, K. M. (1986). Are foreign exchange forecasts rational? New evidence from survey data. *Economics Letters*, 21, 277-281.
- Domowitz, I., & Hakkio, C. S. (1985). Conditional variance and the risk premium in the foreign exchange market. *Journal of International Economics*, 19, 47-66.
- Dumas, B. (1992). Dynamic equilibrium and the real exchange rate in a spatially separated world. *Review of Financial Studies*, 5, 153-180.
- Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica*, 64(4), 813-836.
- Enders, W., & Granger, C. W. J. (1998). Unit root test and asymmetric adjustment with an example using the term structure of interest rates. *Journal of Business & Economic Statistics*, 16(3), 304-311.
- Engel, C. (1992). On the foreign exchange risk premium in a general equilibrium model. *Journal of International Economics*, 32, 305-319
- Engel, C. (1996). The forward discount anomaly and the risk premium: A survey of recent evidence. *Journal of Empirical Finance*, 3(2), 123-192.
- Engel, C. (1999). On the foreign exchange risk premium in sticky-price general equilibrium models. *International Tax and Public Finance*, 6, 491-505.
- Engel, C., & Hamilton, J. (1990). Long swings in the Dollar: Are they in the data and do markets know it? *American Economic Review*, 80, 689-713.
- Engel, C., & Rodrigues, A. P. (1989). Test of international CAPM with time-varying covariances. *Journal of Applied Econometrics*, 4, 119-138.
- Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica*, 55(2), 251-276
- Evans, G. W. (1986). A test for speculative bubbles in the Sterling Dollar exchange rate: 1981-84. *American Economic Review*, 76, 621-636.
- Evans, M. D., & Lewis, K. K. (1993). Trends in excess returns in currency and bond markets. *European Economic Review*(37), 1005-1019.

- Evans, M. D., & Lewis, K. K. (1995). Do long-term swings in the dollar affect estimates of the risk premia? *The Review of Financial Studies*, 8(3), 709-742.
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2), 383-417.
- Fama, E. F. (1984). Forward and spot exchange rate. *Journal of Monetary Economics*, 14(3), 319-38.
- Fama, E. F. (1991). Efficient capital markets: II. *Journal of Finance*, 46(5), 1575-1617.
- Fama, E. F., Fisher, L., Jensen, M., & Roll, R. (1969). The adjustment of stock prices to new information. *International economic review*, 10(1), 1-21.
- Flood, R. P., & Garber, P. M. (1980). Market fundamentals versus price-level bubbles: The first tests. *Journal of Political Economy*, 88, 745-770.
- Flood, R. P., & Hodrick, R. J. (1990). On testing for speculative bubbles. *Journal of Economic Perspectives*, 4, 85-101.
- Frankel, J. A. (1973). Elasticity and the interest parity theorem. *Journal of Political Economy*, 81, 741-747.
- Frankel, J. A. (1978). Purchasing power parity: Evidence from the 1920's. *Journal of International Economics*, 8(2), 161-191.
- Frankel, J. A. (1982). A test of perfect substitutability in the foreign exchange market. *Southern Economic Journal*, 49, 406-416.
- Frankel, J. A. (1988). Recent estimates of time-variation in the conditional variance and in the exchange risk premium. *Journal of International Money and Finance*, 7, 115-125.
- Frankel, J. A., & Froot, K. A. (1987). Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review*, 77, 133-153.
- Frankel, J. A., & Poonawala, J. (2010). The forward market in emerging currencies: Less biased than in major currencies. *Journal of International Money and Finance*, 29(3), 585-598.
- Froot, K. A. (1990). Short rates and expected asset returns. *National Bureau of Economic Research*, No. w3247.
- Froot, K. A., & Frankel, J. A. (1989). Forward discount bias: Is it an exchange risk premium? *Quarterly Journal of Economics*, 104(1), 139-161.
- Froot, K. A., & Thaler, R. H. (1990). Anomalies: Foreign exchange. *The Journal of Economic Perspectives*, 4(3), 179-192.
- Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4), 221-238.

- Giovannini, A., & Jorion, P. (1987). Interest rates and risk premia in the stock market and in the foreign exchange market. *Journal of International Money and Finance*, 6, 107-123.
- Giovannini, A., & Jorion, P. (1989). The time variation of risk and return in the foreign exchange and stock markets. *Journal of Finance*, 44, 307-325.
- Goodhart, C., McMahon, P., & Ngama, Y. (1997). Why does the spot-forward discount fail to predict changes in future spot rates? *International Journal of Finance and Economics*, 2(2), 121-129
- Granger, C. W. J. (1999). Aspects of research strategies for time series analysis. New Haven: Conference on New Development in Time Series Economics.
- Granger, C. W. J., & Hyung, N. (1999). Occasional structural breaks and long memory. *Journal of Empirical Finance*, 11, 399-421.
- Granger, C. W. J., & Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *Journal of Empirical Finance*, 11, 399-421.
- Granger, C. W. J., & Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1), 15-29.
- Granger, C. W. J., & Terasvirta, T. (1993). *Modelling non-linear economic relationships*. Oxford: Oxford Univ. Press.
- Hansen, B. E. (1999). The grid bootstrap and the autoregressive model. *Review of Economics and Statistics*, 81, 594-607.
- Hansen, B. E. (2001). The new econometrics of structural change: dating breaks in US labor productivity. *Journal of Economic Perspective*, 15, 117-128.
- Hansen, L. P., & Hodrick, R. J. (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy*, 88, 829-853.
- Harvey, D.I., Leybourne, S.J., & Taylor, A.M.R. (2006). Modified tests for a change in persistence. *Journal of Econometrics*, 134, 441-469
- Hassler, U., & Meller, B. (2014). Detecting multiple breaks in long memory the case of U.S. inflation. *Empirical Economics*, 46(2), 653-680
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46, 1251-1271.
- Hodrick, R. J. (1984). An investigation of risk and return in forward foreign exchange. *Journal of International Money and Finance*, 3, 5-29.
- Hodrick, R. J., & Srivastava, S. (1986). The covariation of risk premiums and expected future spot rates. *Journal of International Money and Finance*, 3, 5-30.

- Horvath, M. T., & Watson, M. W. (1995). Testing for cointegration when some of the cointegrating vectors are prespecified. *Econometric theory*(11), 984-1014.
- Hosking, J. R. (1981). Fractional differencing. *Biometrika*, 68(1), 165-176.
- Huang, R. D. (1981). The monetary approach to exchange rate in an efficient foreign exchange market: Test based on volatility. *Journal of Finance*, 36, 31-41.
- Kaminsky, G. L. (1993). Is there a Peso Problem? Evidence from the Dollar/Pound exchange rate, 1976-1987. *American Economic Review* , 83, 450-472.
- Kapetanios, G., & Shin, Y. (2011). Testing the null hypothesis of nonstationary long memory against the alternative hypothesis of a nonlinear ergodic model. *Econometric Reviews*, 30(6), 620-645.
- Kaserman, D. L. (1973). The forward rate: Its determination and behavior as a predictor of the future spot rate. *Proceedings of the American Statistical Association* , (hlm. 417-22).
- Kellard, N., & Sarantis, N. (2008). Can exchange rate volatility explain persistence in the forward premium? *Journal of Empirical Finance*, 15, 714-728.
- Kilian, L., & Taylor, M. P. (2003). Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics*, 60(1), 85-107.
- Kim, C. S., & Phillips, P. C. B. (2000). Modified log periodogram regression. *Working Paper, Yale University*. Yale University, mimeographed.
- Kim, C. S., & Phillips, P. C. B. (2006). Log periodogram regression: The nonstationary case. Yale University: Cowles Foundation for Research in Economics.
- Kim, J. (2000). Detection of change in persistence of a linear time series. *Journal of Econometrics*, 95, 97-116
- King, A. (2011). A comment on: "The solution to the forward-bias puzzle". *Journal of International Financial Markets, Institutions and Money*, 21, 623-628.
- Krasker, W. S. (1980). The "Peso Problem" in testing the efficiency of forward exchange markets. *Journal of Monetary Economics*, 6, 269-276.
- Kwiatkowski, D., Phillips, P. C. B., Shin, Y., & Schdmit, P. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 159-78.
- Ledolter, J. (2009). Estimation bias in the first-order autoregressive model and its impact on predictions and prediction intervals. *Communication in Statistics - Simulation and Computation*, 38(4), 771-787
- LeRoy, S. F. (1984). Efficiency and the variability of asset prices. *American Economic Review*, 74, 183-187.

- Lewis, K. K. (1988). The persistence of the "Peso Problem" when policy is noisy. *Journal of International Money and Finance*, 7, 5-21.
- Lewis, K. K. (1989a). Can learning affect exchange-rate behavior? The case of the Dollar in the early 1980's. *Journal of Monetary Economics*, 23, 79-100.
- Lewis, K. K. (1989b). Changing beliefs and systematic rational forecast errors with evidence from foreign exchange. *American Economic Reviews*, 79(4), 621-36.
- Lewis, K. K. (1991). Was there a "Peso Problem" in the U.S. term structure of interest rates: 1979-1982? *International Economic Review*, 32, 159-173.
- Lewis, K. K. (1995). Puzzles in international financial market. Dalam G. M. Grossman, & K. Rogoff (Ed.), *Handbook of International Economics Volume III* (hlm. 1913-1971). Amsterdam: North-Holland.
- Leybourne, S.J., Kim, T. H., & Taylor, A. M. R. (2007). Detecting multiple changes in persistence. *Studies in Nonlinear Dynamics & Econometrics*, 11(3), 1-32.
- Leybourne, S.J., Taylor, A. M. R., & Kim, T. H. (2006). CUSUM of squares-based tests for a change in persistence. *Journal of Time Series Analysis*, 28(3), 408-433
- Liu, P. C., & Maddala, G. S. (1992). Using survey data to test market efficiency in the foreign exchange markets. *Empirical Economics*, 17, 303-314.
- Liu, W., & Maynard, A. (2005). Testing forward rate unbiasedness allowing for persistent regressors. *Journal of Empirical Finance*, 12, 613-628.
- Lizondo, J. S. (1983). Foreign exchange futures prices under fixed exchange rates. *Journal of International Economics*, 14, 69-84.
- Lo, A. W. (2004). The adaptive markets hypothesis. *Journal of Portfolio Management*, 30(5), 15-29.
- Lo, A. W., & MacKinley, A. C. (2001). *A non-random walk down wall street*. Princeton: Princeton University Press.
- Luintel, K. B., & Paudyal, K. (1998). Common stochastic trends between forward and spot exchange rates. *Journal of International Money and Finance*, 17, 279-297.
- Lyons, R. K. (2001). *The Microstructure Approach to Exchange Rates*. London: MIT Press.
- MacDonald, R., & Taylor, M. P. (1993). The monetary approach to the exchange rate: Rational expectations, long -run equilibrium, and forecasting. *International Monetary Fund Staff Papers*, 40, 89-107.
- Maddala, G. S., & Kim, I. M. (1998). *Unit roots, cointegration and structural change*. Cambridge, UK: Cambridge University Press.

- Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, 17(1), 59-82.
- Mankiw, N., & Shapiro, M. (1986). Do we reject too often? Small sample properties of tests of rational expectations models. *Economics Letters*, 20, 139-145
- Mariott, F.H.C., & Pope, J. A. (1954). Bias in the estimation of autocorrelations. *Biometrika*, 41, 390-402
- Mark, N., Yangru, W., & Weihe, H. (1997). Understanding spot and forward exchange rate regressions. *Journal of Applied Econometrics*, 12(6), 715-734.
- Maynard, A. (2003). Testing for forward rate unbiasedness: On regression in levels and in returns. *Review of Economics and Statistics*, 85(2), 313-327
- Maynard, A. (2006). The forward premium anomaly: Statistical artifact or economic puzzle? New evidence from robust tests. *Canadian Journal of Economics*, 39(4), 1244-1281.
- Maynard, A., & Phillips, P. C. B. (2001). Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics*, 16(6), 671-708.
- Maynard, A., & Shimotsu, K. (2009). Covariance-based orthogonality tests for regressors with unknown persistence. *Econometric Theory*, 25(1), 63-116
- McCallum, B. T. (1994). A reconsideration of the uncovered interest parity relationship. *Journal of Monetary Economics*, 33(1), 105-132.
- Meese, R. A. (1986). Testing for bubbles in exchange rate markets: A case of sparkling rates. *Journal of Political Economy*, 94, 345-373.
- Meese, R. A., & Singleton, K. J. (1982). On unit roots and the empirical modeling of exchange rates. *The Journal of Finance*, 37(4), 1029-1035.
- Mikusheva, A. (2007). Uniform inference in autoregressive models. *Econometrica*, 75(5), 1411-1452.
- Muller, C. (2011). The forward-bias puzzle: Still unsolved. *Journal of International Financial Markets, Institutions & Money*, 21, 605-610.
- Mussa, M. (1982). A model of exchange rate dynamics. *Journal of Political Economy*, 90(1), 74-104.
- Newbold, P., Wohar, M., Rayner, T., Kellard, N., & Ennew, C. (1998). Two puzzles in the analysis of foreign exchange market inefficiency. *International Review of Financial Analysis*, 7(2), 95-111
- Ng, S., & Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica*, 69, 1519-1554.

- Obstfeld, M., & Taylor, A. M. (1997). Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited. *Journal of the Japanese and International Economics*, 11(4), 441-479.
- Ohanissian, A., Russell, J. R., & Tsay, R. S. (2008). True or spurious long memory? A new test. *Journal of Business and Economic Statistics*, 26(2), 161-175.
- Panos, M., Nobay, A. R., & Peel, D. A. (1997). Transaction costs and nonlinear adjustment in real exchange rates: An empirical investigation. *Journal of Political Economy*, 105(4), 862-879.
- Pantula, S. G., & Fuller, W. A. (1985). Mean estimation bias in least squares estimation of autoregressive process. *Journal of Econometrics*, 27, 99-121
- Perron, P. (2006). Dealing with structural breaks. *Handbook of Econometrics: Econometric Theory*. Palgrave Handbook.
- Perron, P., & Qu, Z. (2010). Long-memory and level shifts in the volatility of stock market return indices. *Journal of Business and Economic Statistics*, 28, 275-290.
- Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica*, 55(2), 277-301.
- Phillips, P. C. B. (2007). Unit root log periodogram regression. *Journal of Econometrics*, 138, 104-124.
- Phillips, P. C. B., & Perron, P. (1998). Testing for a unit root in time series regression. *Biometrika*, 75, 335-346
- Pippenger, J. (2011). The solution to the forward-bias puzzle. *Journal of International Financial Markets, Institutions & Money*, 21, 296-304.
- Qu, Z. (2011). A test against spurious long memory. *Journal of Business & Economic Statistics*, 29(3), 423-437
- Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, 23, 1630-1661
- Robinson, P. M. & Henry, M. (1999). Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels. *Econometric Theory*, 15, 299-336
- Rogoff, K. (1979). Expectations and exchange rate volatility. Massachusetts Institute of Technology.
- Roll, R. & Yan, S. (2000). An explanation of the forward premium 'puzzle'. *European Financial Management* 6(2). 121-148
- Romano, J. P., & Wolf, M. (2001). Subsampling intervals in autoregressive models with linear time trends. *Econometrica*, 69, 1283-1314.

- Sakoulis, G., & Zivot, E. (2001). Time variation and structural change in the forward discount: Implications for the forward rate unbiasedness hypothesis. *Working paper, Department of Economics*. University of Washington.
- Sakoulis, G., Zivot, E., & Choi, K. (2010). Structural change in the forward discount: Implications for the forward rate unbiasedness hypothesis. *Journal of Empirical Finance*, 17(5), 957-966.
- Sarantis, N. (1999). Modelling non-linearities in real effective exchange rate. *Journal of International Money and Finance*, 18(1), 27-45.
- Sarno, L. (2005). Viewpoint: Towards a solution to the puzzles in exchange rate economics: where do we stand? *Canadian Journal of Economics*, 38(3), 673-708.
- Sarno, L., Taylor, M. P., & Chowdhury, I. (2004). Nonlinear dynamics in the deviations from the law of one price: A broad based empirical study. *Journal of International Money and Finance*, 23, 1-25.
- Sarno, L., Taylor, M. P., & Frankel, J. A. (2003). *The economics of exchange rates*. Cambridge University Press.
- Sarno, L., Valente, G., & Leon, H. (2006). Nonlinearity in deviations from uncovered interest parity: An explanation of the forward bias puzzle. *Review of Finance*, 10(3), 443-482.
- Sercu, P., Vandebroek, M., & Wu, X. (2008). Is the forward bias economically small? Evidence from European rates. *Journal of International Money and Finance*, 27, 1284-1302.
- Sercu, P., & Vinaimont, T. (2006). The forward bias in the ECU: Peso risks vs. fads and fashions. *Journal of Banking and Finance*, 30(8), 2409-2432.
- Sewell, M. (2011). History of the efficient market hypothesis. UCL Department of Computer Science.
- Shaman, P., & Stine, P. (1988). The bias of autoregressive coefficient estimators. *Journal of The American Statistical Association*, 83(403), 842-848.
- Shao, X. (2011). A simple test of changes in mean in the possible presence of long-range dependence. *Journal of Time Series Analysis*, 32(6), 598-606
- Shao, X. & Wu, W.B. (2007). Local whittle estimation of fractional integration for nonlinear processes. *Econometric Theory*, 23, 899-951
- Sharpe, I. (1985). Interest parity, monetary policy, and the volatility of Australian short-term interest rates: 1978-82. *Economic Record*, 61, 436-444.
- Shimotsu, K. (2006). Simple (but effective) tests of long memory versus structural breaks. *Working Paper No. 1101*, Department of Economics, Queen's University

- Shimotsu, K., & Phillips, P. C. B (2006). Local whittle estimation of fractional integration and some of its variants. *Journal of Econometrics*, 103(2), 209-233
- Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics*, 53(1), 165-188.
- Stock, J. H. (1991). Confidence interval for the largest autoregressive root in US macroeconomic time series. *Journal of Monetary Economics*, 29, 435-459.
- Stock, J. H., & Watson, M. W. (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics*, 14(3), 11-30.
- Stockman, A. C. (1978). Risk, information and forward exchange rates. Dalam J. A. Frenkel, & H. G. Johnson, *In the economics of exchange rates: Selected studies*. Boston: Addison-Wesley.
- Takagi, S. (1991). Exchange rate expectations: A survey of surveys studies. *International Monetary Funds Staff Papers*, 38(1), 156-83.
- Taqqu, M.S., & Teverovsky, V. (1996). Semi-parametric graphical estimation techniques for long-memory data. *Athens Conference on Applied Probability and Time Series, Vol. II*, New York: Springer, 420-432
- Tauchen, G. (2001). The bias of tests for a risk premium in forward exchange rates. *Journal of Empirical Finance*, 8, 695-704.
- Taylor, A. M. R. (2001). Potential pitfalls for the PPP puzzle? Sampling and specifications biases in mean reversion tests of the LOOP. *Econometrica*, 69, 473-498.
- Taylor, A. M. R. (2005). Fluctuation tests for a change in persistence. *Oxford Bulletin of Economics and Statistics*, 67, 207-230
- Taylor, M. P. (1988). What do investment managers know? An empirical study of practitioners' predictions. *Economica*, 55, 185-202.
- Taylor, M. P. (1995). The economics of exchange rates. *Journal of Economic Literature*, 33, 13-47.
- Taylor, M. P., & Peel, D. A. (2000). Nonlinear adjustment, long-run equilibrium and exchange rate fundamentals . *Journal of International Money and Finance*, 19, 33-53.
- Taylor, M. P., Peel, D. A., & 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.
- Terasvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of American Statistics Association*, 89, 208-218.

- Topbas, N. (2014). Tests of rationality in Turkish foreign exchange market. *Central Bank Review*, 14(2), 65-78.
- Torous, W., Valkanov, R., & Yan, S. (2004). On predicting stock returns with nearly integrated explanatory variables. *Journal of Business*, 77(4), 937-966
- Velasco, C. (1999). Non-stationary log-periodogram regression. *Journal of Econometrics*, 91(2), 325-371.
- Wang, P., & Wang, P. (2009). Does a "correct" parameter estimate tell a better story about foreign exchange market efficiency? *Journal of International Money and Finance*, 28(2), 183-197.
- Zeileis, A. & Kleiber, C. (2005). Validating multiple structural change models: A case study. *Journal of Applied Econometrics*, 20(5), 685-690.

University of Malaysia