Research Article   **OMICS International**

Practice-Oriented Model Selection in Forex Market

Li Dai[1] and Peng Zhou[2]*

[1]School of Economics and Trade, Hunan University, Fenglin Road, Changsha, 410079, PR China
[2]Cardiff University, UK

Abstract

Most of the theoretical and empirical literatures are dedicated to selecting a more empirically sound model in economics and finance. When choosing from competing alternatives, the most popular criterion is econometric goodness of fit of models against structured data (cross-sectional, time-series and panel). It is the data that determines whether a theory is valid or not, but in many scenarios, structured data are not available or not precise measures of the reality. Therefore, this paper attempts to propose a complementary criterion in model selection in terms of investment practice—we are not trying to find an econometrically "true" model, but a practically useful model. A series of models are examined by the latest data in the FOREX market. Monetary CAPM turns out to be the "best" model for practical purpose.

Keywords: Forex market; Model selection; Practice oriented criterion; Multivariate-Garch-in-Mean; Portfolio management

Introduction

Foreign exchange (FOREX) is required for transactions on current account (goods and services) and capital account (financial assets). In FOREX market, exchange rate is the key variable linking all prices in different countries. Traditionally, it is assumed that exchange rate is determined fundamentally by current account transactions in a long-term perspective (PPP) and by capital account transactions in a short-term perspective (UIP and CIP). As the widespread removal of capital controls, the capital account transactions become the main determinant for the exchange rate. Econometric techniques are usually used as the model comparison criterion, such as likelihood ratio test or Bayes factor. However, these techniques rely on the validity and availability of structured and continuous data measuring the variables in the model. If not, these techniques are not applicable. Moreover, these econometric techniques are only comparing models with models, rather than comparing models with data directly. After all, one can only conclude which model is more relatively true than others, rather than in an absolute sense. Some recent econometric literature attempts to directly confront models with data using Indirect Inference [1,2], but again the selection criterion depends on the choice of auxiliary model through which the model features and the data features are compared. Seeing the limitation of econometric criterion in model selection, this paper proposes a new criterion with less ambition. We are not trying to find the "true" model. Rather, we are looking for the most useful model—the *practice-oriented criterion*—to generate the highest return in investment practice.

As shown in the Flow chart 1, the process of model selection can be divided into 3 stages, through which the unsuitable candidate models are eliminated. Stage 1 is to compare the long-term model and the short-term model. Stage 2 is to compare the risk neutral model and the risk adjusted model. Stage 3 is finally to compare different specifications of risk premium to find the best model. Sophisticated econometric methods are involved in this process.

Following this Flow chart 1, section 2 compares the prevailing models in the theoretical and empirical finance literature to select the best model in terms of investment practice, the conclusion of which is then applied in section 3 with a case study in FOREX market.

Model Selection

Stage 1: long-term model *vs.* short-term model

As the benchmark of long-term models, Purchasing Power Parity (PPP) is a simple proposition that "once converted to a common currency, national price levels should be equal", so that a unit of currency of one country will have the same purchasing power in a foreign country. The underlying principle is that *arbitrage* opportunities

Flow chart 1: Process of model selection.
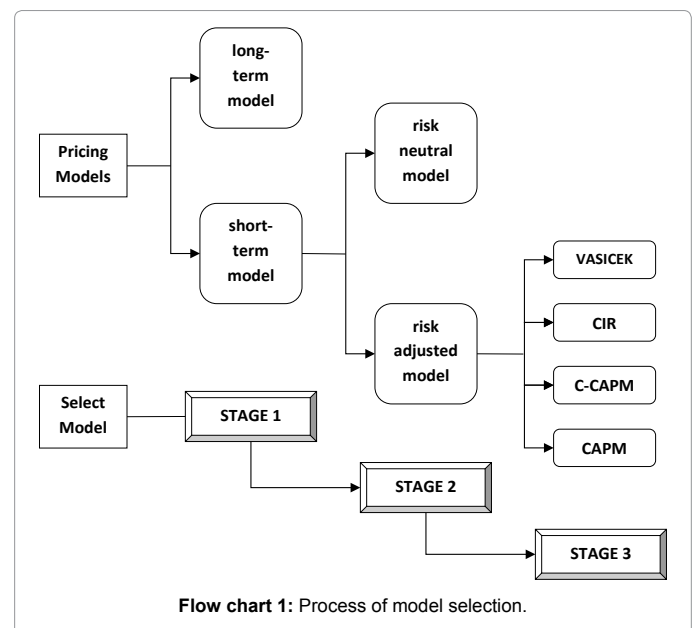
***Corresponding author:** Peng Zhou, Lecturer in Economics, D47, Aberconway Building, Colum Drive, Cardiff University, CF10 3EU, UK, Tel: 0044-7528223250; E-mail: pzhou@cardiffmet.ac.uk

should be eliminated in the long run. There are various versions of this famous theory, so it is necessary to review some of the variants used in both theory and practice when judging the validity of this theory.

The law of one price: The basic building block for any variation of PPP is the "law of one price" (LOP). It states that for any good i= $P_{i,t} = S_t \cdot P_{i,t}^*$, where $P_{i,t}$ is the domestic price of good i at t, $P_{i,t}^*$ is the foreign price, and S_t is the spot nominal exchange rate, defined as the domestic currency price of foreign currency. In other words, LOP claims that the translated prices for any identical good in different countries should be the same in an efficient market.

Obviously, LOP is unlikely to hold due to natural costs (such as transportation costs), artificial costs (such as tariff or non-tariff barriers), and other economic factors (such as asymmetric information and competition). One of the most famous tests of LOP is the prices of McDonald's "Big Mac" hamburger across the world. According to the survey of *The Economist* newspaper on 15th April 1995, the prices for the same "Big Mac" in different countries ranged from $5.20 in Switzerland at the highest end to $1.05 in China at the lowest end [3]. In January 2004, the cheapest burger was still in China, at $1.23, compared with an average American price of $2.80 [4]. The conclusion is obvious that LOP is rejected. However, LOP is not designed for practice, but for theoretical need.

Absolute PPP: Based on LOP, absolute PPP, a slightly more complicated theory, is developed. It requires that: $P_t = S_t \cdot P_t^*$, where P_t and P_t^* are sums taken over the weighted average prices to construct the price indices (using Consumer Price Index or Producer Price Index). The focus is then diverted from a *single* good to the *overall* price level, compared to the LOP. However, the problems are exacerbated as one must worry about how to handle the introduction of new goods, shifting consumption weights within a country, and so on. These new difficulties enfeeble absolute version of PPP.

Relative PPP: Relative version of PPP only requires that the rate of growth in the exchange rate should offset the differential between the rates of growth in domestic and foreign price indices: $(P_t / P_{t-1}) = (S_t / S_{t-1}) \cdot (P_t^* / P_{t-1}^*)$. Compared with absolute PPP, this equation diverts the focus from *static* to *dynamic*.

Relationship: The relationship between absolute and relative versions of PPP is direct. "If absolute PPP holds, then relative PPP must also hold. However, if relative PPP holds, then absolute PPP does not necessarily hold." We can examine this proposition from a broader perspective. Rewrite the equation in the form of $(S_{t-1} \cdot P_{t-1}^* / P_{t-1}) = S_t \cdot P_t^* / P_t$, and we can define a new concept "*Real Exchange Rate*" as $Q_t = S_t \cdot P_t^* / P_t$. Compared to this general notion, both versions of PPP are actually special cases, presuming that real exchange rate is constant throughout time. The absolute PPP assumes $Q_t = Q_{t-1} = 1$ and the relative PPP assumes $Q_t = Q_{t-1} =$ any arbitrary constant. Hence, relative PPP includes absolute PPP as a special case. In the form of natural logarithm, the relative PPP can be written as $q_t = s_t + p_t^* - p_t = k$, where k=0 for the absolute PPP case. The lowercase denotes the logarithm of the uppercase and *k* is a constant. In the light of this relation, one did not have to be an econometrician to witness the "collapse of purchasing power parity": one could simply examine the behavior of the real exchange rate. That is to say, to reject PPP, usually relative PPP, is equivalent to reject the constant real exchange rate.

Model: There have been tons of literatures on testing PPP, but most models are based on this regression: $q_t = \alpha + \beta \cdot q_{t-1} + \varepsilon_t$, where ε_t denotes the random disturbance. If β=1, we say that the real exchange rate is

subject to a *unit root process*, which does not revert to any average level over time. The unit root process property of real exchange rate implies that shocks never die out in the long run because it has no tendency or mean reversion. Thus, the test for null hypothesis H_0: β=1, is a test for whether long run PPP does not hold. According to an early study in this spirit, using annual data from 1869 to 1984 for the dollar-sterling real exchange rate, Frankel [5,6] estimates a first order autoregressive process for the real exchange rate in the form:

$$q_t - \tilde{q} = \varphi(q_{t-1} - \tilde{q}) + \varepsilon_t$$

$\tilde{q}$ is the assumed constant equilibrium level of real exchange rate, and φ is the autocorrelation coefficient - an unknown parameter describes the speed of mean reversion. A proportion of of any shock will still remain after one period, [2] of it remain after two periods, and in general, φ^n of the shock will remain after n periods. Intuitively, we can consider the speed of adjustment by asking how long it would take for the effect of a shock to die out by 50%, i.e. computing the half-life of shocks.

Evidence: Frankel's estimate of φ is 0.86, which implies a half-life of about 5 years.[1] Similar results were found by Edison (1987), based on the data over the period 1890-1978, and by Glen, using the data sample spanning the period 1900 - 1987. Lothian and Taylor [4] used two centuries of data on dollar-sterling and franc-sterling real exchange rates, reject the random walk hypothesis and find point estimates of φ of 0.89 for dollar-sterling and of 0.76 for franc-sterling. Obstfeld and Rogoff [3,7] used the data between the period 1973 - 1995, obtaining a value of 0.99 between the U.S. and Canadian dollars, which implies a half-life of 69 months.

Since the latest data up to 2016 is available, a new test can be carried out to test the reliability of PPP in FOREX market during the latest decades. I use the monthly data between 1960 and 2016[2] to run an OLS regression based on the autoregressive model: $q_t = \alpha + \beta \cdot q_{t-1} + \varepsilon_t$. The results are as follows[3]: the estimated rate of adjustment β is 0.99, which is not significantly different from 1 at 5% level, similar to the results of Obstfeld and Rogoff. In other words, the estimated half-life of the shock is around 69 months, i.e. more than 5.5 years. This result implies the nonstationarity of real exchange rate, which leads to a troublesome riddle which we call PPP Puzzle (Figure 1 and Table 1).

PPP puzzle: Based on the observation in hundreds of studies using widely varying techniques and data sets, researchers have repeatedly found very long half-lives on the order of 3 to 5 years, for shocks to real exchange rates. As stated by Rogoff, "The purchasing power parity puzzle then is this: How can one reconcile the enormous short-term volatility of real exchange rates with the extremely slow rate at which shocks appear to damp out?" In the test using the 2016 data, I also find a sluggish speed of adjustment of 62 months. Hence, the puzzle remains in so far as the latest data tells us.

Explanation: The conditions for relative PPP to hold are still demanding. The main criticisms lie in *tradability* of goods and the

---

[1]You could find the way of obtaining the half-life of a shock from any science textbooks and it is not the problem to be solved in this paper. The formula is given without proof here:

$$t_{1/2} = \frac{\ln 2}{1 - \varphi}$$

[2]The time series data of dollar-sterling exchange rates, US CPI and UK CPI are obtained from *DATASTREAM*.

[3]To get an intuitive report, please refer to Figure 1 for the plot of the regression attached in Appendices.
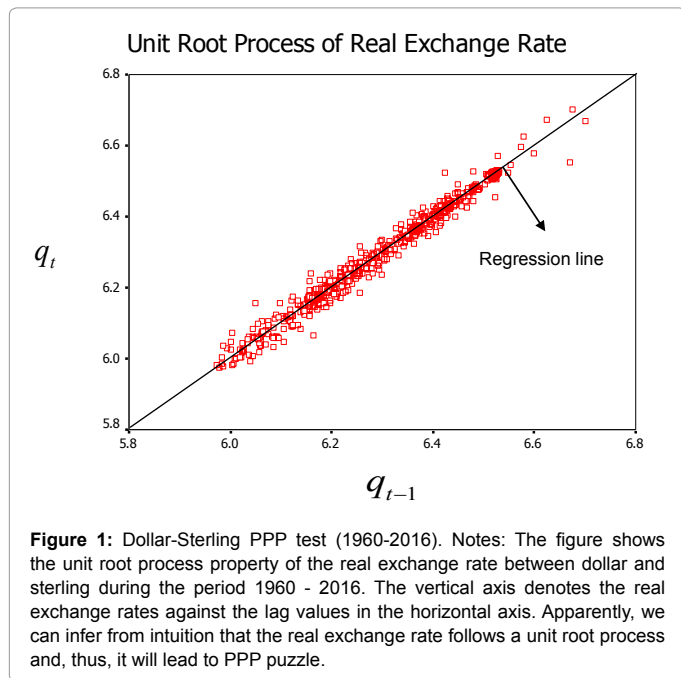
Figure 1: Dollar-Sterling PPP test (1960-2016). Notes: The figure shows the unit root process property of the real exchange rate between dollar and sterling during the period 1960 - 2016. The vertical axis denotes the real exchange rates against the lag values in the horizontal axis. Apparently, we can infer from intuition that the real exchange rate follows a unit root process and, thus, it will lead to PPP puzzle.

Model Summary	
R-Square	0.976541
Adjusted R-Square	0.976498
Observation	558

ANOVA	df	SS	MS	F	Significance F
Analysis of regression	1	11.79872	11.79872	23144.58	0
Residual	556	0.283439	0.00051		
Sum	557	12.08216			

Variable	Coefficients	SE	t-Stat	Lower 95%	Upper 95%
Constant	0.067874	0.04095	1.657496	0.097983	-0.01256
q_{t-1}	**0.989102**	0.006502	152.1334	0	0.976332

Table 1: Summary of level test for Figure 1.

conformability of the basket of goods. *First of all*, it is obvious that not all goods and services are tradable internationally due to considerable transaction costs and/or artificial barriers such as monopolistic power and authoritative tariffs. *Secondly*, as we could expect, the consumption structures are also diverse in different countries due to various cultures. For example, bread is the main food for the western world whilst the Asians usually eat rice. *Thirdly*, Balassa-Samuelson effect[4] can also be utilized to elucidate the puzzle [8]. They argued that, when all countries' price levels are translated into US dollars at the prevailing nominal exchange rate, the richer countries tend to have higher price levels than those of poorer ones. The reason consists in the difference in relative productivity of tradable goods between countries, which exacerbates the differential between postulated and estimated exchange rate.

More fundamentally speaking, the first factor is related to the process of the international economic interflow. The second factor is related to the demand aspect, while the third factor is related to the

[4]**Balassa, B.** (1964) "The Purchasing Power Parity Doctrine: A Reappraisal." *Journal of Political Economy*, vol. 72, December, pp. 584-596. and **Samuelson, P.** (1964) "Theoretical Notes on Trade Problems." *Review of Economics and Statistics*, volume 23. The two famous papers were independently developed but in the same year.

supply aspect. These three factors all contribute to the impedance of the process of *arbitrage*, so that PPP is always not a practical proposition in reality. However, we are not asserting that PPP is valueless. Actually, theoretical delicate models, such as earlier "Monetary Models"[5], later "Sticky Price Models"[6], and "Nonlinear Models"[7], are all based on PPP long run equilibrium. We will resort to the "Monetary Model" in the CAPM model later on. Nevertheless, as this paper is aimed to practice, a sluggish half-life of 3~5 years is too long a horizon for investors to make short run decision. As a result, PPP is not in the short list.

Stage 2: Risk neutral model V. S. risk adjusted model

Compared to PPP, interest rate parity (IRP) is focused on the short-term equilibrium, which is more suitable to be applied for practical use. Just like PPP, IRP theory also has several versions. To further our model selection, it is necessary to begin with specification of these different versions.

Uncovered interest parity (UIP): The underlying principle of this short-term condition is the same as PPP, i.e. no arbitrage condition. The focus, however, is diverted from goods and service markets to financial markets arbitrage opportunities, in other words, from long-term to short-term perspective. Under this principle along with a simplified assumption of *risk neutral investor*, we can arrive at a UIP condition as follows. In order for the investors to be indifferent between the two investments, the expected pay-offs when expressed in the same currency (domestic or foreign currency) must be the same. The only uncertain variable is S_{t+1}, which can be expected conditional on available information at time t. Denote the risk free rates in domestic and foreign countries by r_t and r_t^*. Thus,

$$(1+r_t) = E_t\left[\frac{S_{t+1}}{S_t}(1+r_t^*)\right] \text{ or equivalently } (1+r_t) = \frac{E_t[S_{t+1}]}{S_t}\cdot(1+r_t^*)$$

Take logs and use the approximation $\ln(1+x) \approx x$ when x is small, we can get: $r_t = r_t^* + E_t[\Delta s_{t+1}]$ or equivalently,

$$E_t[s_{t+1}] = s_t + r_t - r_t^*$$

Where s=lnS. This UIP condition gives the market expected future (log) spot rate as linear function of variables known at time t.

Covered Interest parity (CIP): In any developed FOREX markets, forward rate can be quoted to forecast the spot rate in a future date. To avoid the uncertainty associated with not knowing s_{t+1} at time t, investors typically hedge their portfolio by taking out a forward contract. This fixes the exchange rate at which the foreign bond proceeds at time t+1 will be converted into domestic currency. Let $F_{t,t+1}$ denote the forward exchange rate at date t for delivery at time t+1. Similar to UIP, CIP condition can be written in the form: $f_{t,t+1} = s_t + r_t - r_t^*$, where $f_{t,t+1} = \ln F_{t,t+1}$.

Risk adjusted uncovered interest parity (RAUIP): A crucial implicit assumption of UIP is that investors are risk neutral, which is usually not realistic. As a result, there is no risk premium associated with the risky investment of holding the foreign asset. The risk is due to s_{t+1} being unknown at time t. If, as seems far more likely, given the prevalence of currency hedging, investors are risk averse then they

[5]This approach assumed that the PPP held continuously (Frenkel, 1976; Taylor, 1995; Frankel and Rose, 1995).

[6]This approach assumed that the PPP held discontinuously because of sticky price (Dornbusch, 1976).

[7]This approach assumed that the PPP exchange rate held until the arbitrage profit meets a threshold. It implies a nonlinear speed of adjustment towards the equilibrium.

would require a higher return on the foreign investment to compensate for the FOREX risk. If ρ_t denotes risk premium for holding the foreign asset from t to t+1, then UIP would be replaced by RAUIP:

$$1 + r_t + \rho_t = E_t\left[\frac{S_{t+1}}{S_t}(1 + r_t^*)\right] \text{ or approximately } E_t[s_{t+1}] = s_t + \rho_t + r_t - r_t^*$$

Compared to UIP, to determine the expected future spot exchange rate would need to take into account the risk premium ρ_t in RAUIP. Nevertheless, CIP condition will not change after risk adjustment, because forward rate resolves the uncertainty of the payoffs and the investment will be risk free.

Relationship: If the arbitrage opportunities are strictly eliminated under risk neutral assumption in the FOREX market or, in other words, the FOREX market behaves *unbiasedly*[8], then the forward rate $f_{t,t+1}$ should be an unbiased predictor of s_{t+1}. Combine UIP and CIP, we can get:

$$f_{t,t+1} = E_t[s_{t+1}] = s_t + r_t - r_t^*,\qquad(1)$$

or equivalently, $s_{t+1} = f_{t,t+1} + \varepsilon_{t+1}$, where, $E_t[\varepsilon_{t+1}] = 0$

Then this condition can be seen as the exchange rate determination equation provided that the investors are justified to be risk neutral. Meanwhile, it can also be seen as the null hypothesis for *unbiased market*.

Similarly, if the arbitrage opportunities are strictly eliminated under risk averse assumption or, in other words, the FOREX market behaves *efficiently*, then the relation between forward rate $f_{t,t+1}$ and future spot rate s_{t+1} can be found by combining RAUIP and CIP:

$$f_{t,t+1} + \rho_t = E_t[s_{t+1}] = s_t + \rho_t + r_t - r_t^*,\qquad(2)$$

or equivalently, $s_{t+1} = f_{t,t+1} + \rho_t + \varepsilon_{t+1}$, where $E_t[\varepsilon_{t+1}] = 0$

Then this risk adjusted condition can be seen as alternative exchange rate determination equation provided that the investors are risk averse. Meanwhile, it can also be seen as the null hypothesis for *efficient market*. The divergence between equation (1 and 2) lies in risk attitude of investors, then empirical tests can be carried out to see which one is closer to real world.

Model: It is easier to begin with the test of "unbiased market" proposition equation (1), which can be regarded as the benchmark for equation (2). There are two alternative ways to test the unbiasedness hypothesis: level test and difference test.

Level Test is based on the ability of the forward rate to predict the level of spot rate, with a null hypothesis:

H_0: $s_{t+1} = f_{t,t+1} + \varepsilon_{t+1}$, with $E_t[\varepsilon_{t+1}] = 0$

The first part implies no arbitrage opportunity under risk neutral assumption, and the second part of the hypothesis implies rational expectations, i.e. ε_{t+1} is serially uncorrelated with zero mean. It is common to assume weakly rational expectations when the information set consists of current and past values of exchange rates and forward rates, i.e. $I_t = \{s_t, s_{t-1}, \dots, f_{t,t+1}, f_{t-1,t}\}$. To carry out the level test, we need an alternative hypothesis:

H_1: $s_{t+1} = \alpha + \beta.f_{t,t+1} + e_{t+1}$ (3)

This general econometric model provides the testable restrictions under null hypothesis that =0, β=0, and $E_t[e_{t+1}]=0$. Consequently, the alternative hypothesis holds when any of these conditions is violated.

Difference Test is based on the ability of the forward premium to predict the change of the spot rate, with a null hypothesis:

H_0: $\Delta s_{t+1} = f_{t,t+1} - s_t + {}_{t+1}$, with $E_t[\varepsilon_{t+1}]=0$

The alternative hypothesis can then be written as:

H_1: $\Delta s_{t+1} = \alpha + \beta.(f_{t,t+1} - s_t) + e_{t+1}.$ (4)

The null hypothesis holds only when α=0, β=1 and $E_t[e_{t+1}]=0$

Evidence: Surprisingly, it transpires that these two formulations of the test for FOREX market unbiasedness give very different results. In some level tests, $\hat{\beta} = 1$ but the error term always displays serially correlation. In the difference test model, however, the null hypothesis is entirely rejected. The estimates of β are significantly different from 1 and the error term displays serially correlation. Bilson and Fama [9] document the finding that $\hat{\beta} < 1$. Froot [10] summarizes that the average value of $\hat{\beta}$ in over 75 published estimates is -0.88. McCallum even reports an average value of -4, using monthly data from 1978 to 1990. Only a few of the estimates are greater than zero, and none is greater than or equal to 1. To reconcile the different outcomes between level test and difference test can be, Smith and Wickens [11] point out that it is because "s_{t+1} and $f_{t,t+1}$ are nonstationary processes, but the risk premium is stationary". Hence, "super-consistent estimates are obtained" in level tests.
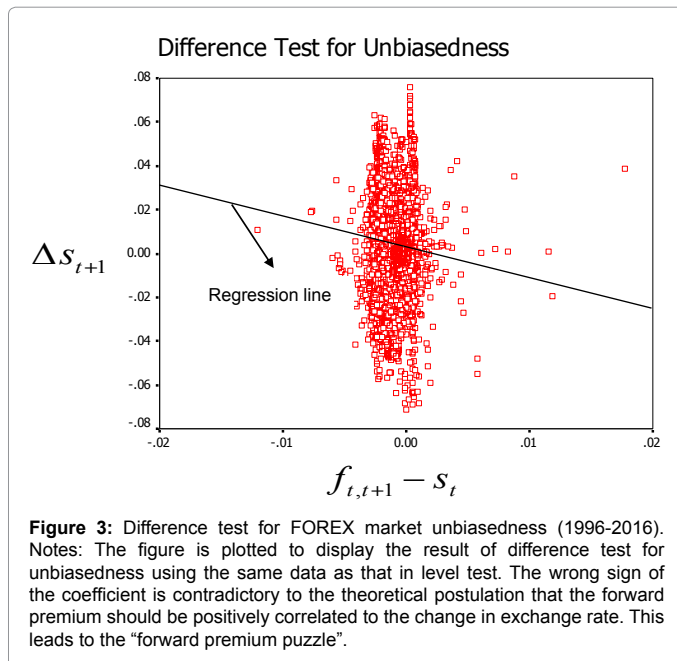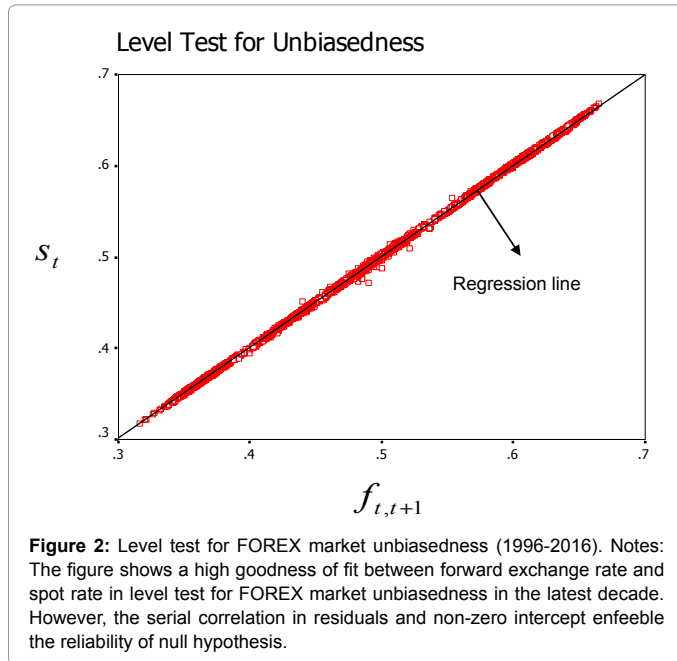
I carry out both the level and difference tests based on empirical models 3 and 4, using the latest dollar-sterling monthly data[9] during 1996-2016. The purpose of this new empirical test is to focus on the last decade and examine the reliability of market unbiasedness hypothesis. As exhibited below, the results are still similar to earlier studies. The estimate of β in level test is 0.97, which is close to 1. Whereas the bad news is that $\hat{\alpha}$ is significantly different from 0 and Durbin-Watson statistic suggests serial correlation in residuals. In contrast, $\hat{\beta}$ in difference test is still negative, -0.93, but $\hat{\alpha}$ is not significantly different from 0. Serial correlation is also found in difference test (Figures 2, 3, Tables 2 and 3).

Forward premium puzzle: According to the empirical evidence, there are some intricate implications in common. Firstly, $(f_{t,t+1} - s_t)$ has the wrong sign and explains little of the variation in Δs_{t+1}. Secondly, both regressions display serially correlated errors, which mean the expectation is irrational. Thirdly, the small R^2 implies that e_{t+1} is the biggest factor affecting Δs_{t+1}. These findings are nominated by researchers as "Forward Premium Puzzle", "Forward Discount Puzzle"[10], or "Predictable Excess Return Puzzle". The forward premium predicts exchange rate change but typically with the wrong sign and smaller magnitude than specified by rational expectations. It is one of the most prominent empirical riddle in international finance.

Explanation: There are several explanations to this puzzle. Basically, they can be classified into two categories. The first type of interpretations focuses on the specification of the model, suggesting adding in a risk premium. The other type of interpretation resorts to various anomalies.

[8]To follow the convention in this domain, "UNBIASED market" means no arbitrage opportunity under risk neutral assumption. In contrast, "EFFICIENT market" means no arbitrage opportunity under risk averse assumption. Hence, the difference between unbiasedness and efficiency is just whether the exchange rate determination equation should include a risk premium term.

[9]The time series data of dollar-sterling exchange rates and forward rates are obtained from *DATASTREAM*.

[10]See Jeffrey Frankel and Jumana Poonawala (2006) "The Forward Market in Emerging Currencies: Less Biased than in Major Currencies".

Figure 2: Level test for FOREX market unbiasedness (1996-2016). Notes: The figure shows a high goodness of fit between forward exchange rate and spot rate in level test for FOREX market unbiasedness in the latest decade. However, the serial correlation in residuals and non-zero intercept enfeeble the reliability of null hypothesis.

Figure 3: Difference test for FOREX market unbiasedness (1996-2016). Notes: The figure is plotted to display the result of difference test for unbiasedness using the same data as that in level test. The wrong sign of the coefficient is contradictory to the theoretical postulation that the forward premium should be positively correlated to the change in exchange rate. This leads to the "forward premium puzzle".

(1) Omitting a risk premium[11]:

If the investor is risk averse, then the UIP condition should be replaced by RAUIP: $E_t\left[\Delta s_{t+1}\right] = \rho_t + r_t - r_t^*$, so the level and difference tests will be modified.

Level test: H_0: $s_{t+1} = \rho_t + f_{t,t+1} + \varepsilon_{t+1} \Leftrightarrow h_1$: $\Delta s_{t+1} = \alpha + \rho_t + \beta \cdot f_{t,t+1} + \varepsilon_{t+1}$

Difference test: H_0: $\Delta s_{t+1} = \rho_t + f_{t,t+1} - s_t + \varepsilon_{t+1} \Leftrightarrow H_1$: $\Delta s_{t+1} = \alpha + \rho_t + \beta \cdot (f_{t,t+1} - s_t) + \varepsilon_{t+1}$

The RAUIP condition implies that $(f_{t,t+1} - s_t)$ should be negatively

---

Model Summary	
R-Square	0.920458
Adjusted R-Square	0.920429
Durbin-Watson	0.094
Observation	2773

ANOVA	df	SS	MS	F	Significance F
Analysis of regression	1	17.38687	17.38687	32065.83	0
Residual	2771	1.502504	0.000542		
Sum	2772	18.88938			

Variable	Coefficients	SE	t-Stat	Lower 95%	Upper 95%
Constant	0.016033	0.002677	5.989611	2.38E-09	0.010784
$f_{t,t+1}$	0.972677	0.005432	179.0693	0	0.962026

Table 2: Summary of level test for Figure 2.

Model Summary	
R-Square	0.003624
Adjusted R-Square	0.003262
Durbin-Watson	0.104
Observation	2752

ANOVA	df	SS	MS	F	Significance F
Analysis of regression	1	0.005017	0.005017	10.00256	0.00158
Residual	2750	1.379404	0.000502		
Sum	2751	1.384422			

Variable	Coefficients	SE	t-Stat	Lower 95%	Upper 95%
Constant	0.000619	0.000522	1.184836	0.236185	-0.00041
$(f_{t,t+1} - s_t)$	-0.93071	0.294278	-3.16268	0.00158	-1.50774

Table 3: Summary of difference test.

correlated with the risk premium *ceteris paribus*. If the risk premium is omitted, the estimate of coefficient of $(f_{t,t+1} - s_t)$ will be biased due to the dependence between disturbance and the regressor. This explanation is consistent with the empirical evidence:

$$Cov[\rho_t, f_t - s_t] < 0 \Rightarrow E[\tilde{\beta}_{OLS}] = \beta + \frac{Cov[\rho_t, f_{t,t+1} - s_t]}{Var[\Delta s_{t+1}]} < \beta$$

It is necessary to notice that actually, there are many literatures rejecting to use risk premium as an interpretation to this puzzle, but no one could provide a better framework to persuade everyone to give up this approach either. At the moment, we still base our model on risk premium approach. Although it might not be the best forever, it is still the most feasible and prevailing asset pricing model for now and for our practical purpose.

(2A) Market anomalies: expectational errors [10]

There is a strand of literature dedicated to expectation errors to interpret the puzzle, including noise trading, peso effect, learning etc. Since the underlying logic of these problems is generic, here we can just consider a representative, noise trading. Suppose that a proportion θ of the investors are using correct model (UIP), and the noise traders assume that there is no change in the exchange rate (e.g. random walk).

$$E_t[s_{t+1}] = \theta \cdot E_t[s_{t+1}|\text{correct model}] + (1-\theta) \cdot E_t[s_{t+1}|\text{wrong model}]$$
$$= \theta \cdot (r_t - r_t^* + s_t) + (1-\theta) \cdot s_t$$
$$= s_t + \theta \cdot (r_t - r_t^*)$$
$$\Rightarrow E_t[\Delta s_{t+1}] = \theta \cdot (r_t - r_t^*) = \theta \cdot (f_t - s_t)$$

Compare this theory with the original regression model, it is

---

obvious that θ=β. However, $\hat{\beta} < 0$ which contradicts the assumption θ > 0. Therefore the argument seems to be problematic in this sense[12].

(2B) *Market anomalies: threshold effects or nonlinearity*[13]

Assume that due to the considerable transaction costs, investors do not response to interest differential until it is large. This proposition implies that UIP only holds for large values of the interest differentials, with a *nonlinear* speed of adjustment. Nevertheless, the empirical evidence shows that both large positive and large negative values of the interest differentials tend to be associated with large negative values of Δs_{t+1}, contradicting the hypothesis that large differentials and large positive values of Δs_{t+1} should move in the same direction. On this evidence, therefore, the threshold hypothesis obtains little support.

(2C) *Market anomalies: carry trade (Burnside)*

Carry trade assumes that the investors search globally only for the highest rate of return regardless of the currency of denomination. Chasing the highest return will cause an increase in demand for that currency, which consequently leads to appreciation in that currency. Thus, instead of interest differential existing to compensate the expected future depreciation, in fact there should be an appreciation.

$$\begin{cases} r_t \uparrow \to demand \uparrow \to E_t[s_{t+1}] \downarrow \\ r_t = r_t^* + (f_t - s_t) \quad (UIP + CIP) \end{cases} \Rightarrow r_t \uparrow \to (f_t - s_t) \uparrow \to E_t[s_{t+1}] \downarrow \to E_t[\Delta s_{t+1}] \downarrow$$

Carry trade is a rational approach to market behavior in the sense that it is self-fulfilling prophecy just like rational bubble. Hence, the exchange rate will be far away from its fundamental value and there will be a sharp reversion back to its fundamental value at some point. This interpretation is consistent with the facts.

To summarize these explanations to this puzzle, risk premium approach and carry trade approach are short-listed in our scenario. Carry trade, however, is more a qualitative interpretation than a quantitative and formulable restriction. Hence, in this round, risk adjusted model beats the risk neutral model and other candidates. The next step is then to specify the risk premium.

Stage 3: Specifications of risk premium

As suggested in Stage 2, the evidence is strongly consistent with the omission of a time-varying risk premium and the no arbitrage condition is:

$$E_t\left[\Delta s_{t+1}\right] = \rho_t + r_t - r_t^*$$

To advance our model in the risk averse paradigm, there are two issues to be settled down first. On one hand, we need to introduce various asset pricing theories to specify the FOREX risk premium. On the other hand, it is necessary to discuss the structure of the FOREX market, since the risk perceived by domestic investors may be different from that perceived by foreign investors due to the currency on which they are based. However, it is only the beginning of the analysis. To arrive at the final conclusion, the two separate parts should be combined to generate a series of testable empirical models. After that, some complicated econometric techniques will be used to estimate the risk premium and examine the statistic properties, resulting in the best model in so far as the latest data tells us.

---

[12]See Mark and Wu (1998), De Long et al. (1990) for more discussion on noise traders.

[13]Asymmetries in foreign exchange markets are supported by evidence of conditional-mean nonlinearities in nominal and real exchange rates (Kräger and Kugler, 1993; Obstfeld and Taylor, 1997; O'Connell, 1998; Taylor, Peel and Sarno, 2001; Coakley and Fuertes, 2001b).

Pricing theory: To specify the risk premium, various asset pricing theories are inevitably resorted to. Among those sophisticated and delicate pricing theories, the Stochastic Discount Factor (SDF) model is the most general and convenient way to price assets. Most existing pricing models can be shown as special cases of SDF model, including Capital Asset Pricing Model (CAPM) of Sharpe-Lintner-Black, Consumption based CAPM (C-CAPM) of Rubinstein and Affine Factor models of Vasicek and Cox-Ingersoll-Ross. Except for SDF model family, the famous Arbitrage Pricing Theory (APT) of Ross is another strand of pricing theory. As suggested by Smith and Wickens, "a key feature of SDF, not possessed by APT, is that the factors in SDF models are linear functions of conditional covariance between the factors and excess return on the risky asset" [12].

This paper will mainly concentrate on SDF models.

(1) Basic SDF model:

The SDF model starts with a very simple proposition that the asset price at period t is the expected discounted value of the asset's payoff in period t+1, conditional on the information available in period t.

$P_t = E_t [M_{t+1}.X_{t+1}]$ or equivalently, $1 = E_t [M_{t+1}.R_{t+1}]$ (5)

P_t is the price or the present value of the asset in period t.

M_{t+1} is the stochastic discount factor for period t+1. ($0 \le M_{t+1} \le 1$)

X_{t+1} is the future payoff of the asset in period t+1.

R_{t+1} is the gross return of the risky asset, equal to $1 + r_{t+1}$.

Mathematically, the equation (5) can be transformed equivalently as follows:

$$1 = E_t[M_{t+1} \cdot R_{t+1}] = E_t[M_{t+1}] \cdot E_t[R_{t+1}] + Cov_t[M_{t+1}, R_{t+1}]$$
$$\Rightarrow E_t[R_{t+1}] = \frac{1 - Cov_t[M_{t+1}, R_{t+1}]}{E_t[M_{t+1}]} \quad (6)$$

Equation (6) holds no matter whether the asset is risk free or risky. As a special case, the gross return of risk free asset is known in period t and can be written in the form of $R_{t+1} = 1 + r_t^f$. Hence, we can utilize this information in (6).

$$E_t[R_{t+1}] = 1 + r_t^f = \frac{1}{E_t[M_{t+1}]} \Rightarrow E_t[M_{t+1}] = \frac{1}{1 + r_t^f} \quad (7)$$

Substitute this result back into equation (6) to get the excess return:

$$E_t[R_{t+1}] = (1 + r_t^f)(1 - Cov_t[M_{t+1}, R_{t+1}])$$
$$\Rightarrow E_t[r_{t+1} - r_t^f] = -(1 + r_t^f) \cdot Cov_t[M_{t+1}, R_{t+1}]$$
$$E_t[r_{t+1} - r_t^f] = -(1 + r_t^f) \cdot Cov_t[M_{t+1}, r_{t+1} - r_t^f] \quad (8)$$

The right hand side of eqn (8) is the risk premium, the extra return over the risk free rate required to compensate the risk averse investor for holding the risky asset. This is also the no arbitrage condition that all correctly priced assets satisfy. Conventionally, the risk premium is broken down in two parts, the price of risk and the quantity of risk.

$$E_t[r_{t+1} - r_t^f] = -(1 + r_t^f) \cdot Cov_t[M_{t+1}, r_{t+1} - r_t^f]$$
$$= \underbrace{\left\{-(1 + r_t^f) \cdot \frac{Cov_t[M_{t+1}, r_{t+1} - r_t^f]}{SD_t[r_{t+1} - r_t^f]}\right\}}_{\text{Price of Risk}} \cdot \underbrace{SD_t[r_{t+1} - r_t^f]}_{\text{Quantity of Risk}}$$
$$= \left(\text{Price of Risk}\right) \times \left(\text{Quantity of Risk}\right)$$

A widely used approximation is due to be introduced here to advance our analysis. The additional assumption is that the stochastic

discount factor and the gross return are jointly lognormally distributed. The *lognormal lemma* tells us that: If ln x is subject t, $N[\mu,\sigma^2]$, then x is lognormally distributed with expectation $E[x]=\exp[\mu + \sigma^2/2]$. Recall another relevant approximation, if x is small, then $\ln(1+x) \approx x$. Based on this new assumption and the two rules, we can redevelop the expression of excess return from equation (5).

$$1=E_t[M_{t+1}. R_{t+1}]=E_t[\exp(\ln M_{t+1}\ln R_{t+1})]$$

$$=E_t[\exp(m_{t+1}+r_{t+1})] \text{ (log approximation)}$$

$$=\exp\{E_t[m_{t+1}]+E_t[r_{t+1}]+Var_t[m_{t+1}+r_{t+1}]\} \text{ (lognormal lemma)}$$

Similar to the earlier procedure, utilize the special case of risk free asset to get a useful restriction, then substitute it back to the original equation, resulting in:

$$E_t[r_{t+1}-r_t^f]+\frac{1}{2}Var_t[r_{t+1}]=-Cov_t[m_{t+1},r_{t+1}] \qquad (9)$$

Equation (9) is the no arbitrage condition of SDF under lognormal assumption, pushing (8) forward. It comprises a term on left hand side, i.e. half conditional variance of return. This is called "*Jensen effect*", which arises because the expectation is taken of a nonlinear function. It is often not included in risk premium since it is comparatively small and ignorable. Hence, the term on the right hand side is the new version of risk premium under lognormal assumption. However, basic SDF model does not provide any testable restriction since m_{t+1}, known as the "pricing kernel", are still a black box. Thus, we need more detailed theories to embody the stochastic discount factor.

(2) C-CAPM:

Consumption-based Capital Asset Pricing Model (C-CAPM) is considered to be a general equilibrium model. Asset pricing is then put into an *intertemporal* optimization problem for a risk averse investor.

$$\max_{C_t}\{U(C_t)+\beta \cdot E_t[U(C_{t+1})]\}, \text{ subject to: } W_{t+1}=W_t \cdot (1+r_{t+1})+P_t \cdot (Y_t-C_t).$$

U (.) is a concave utility function of nominal consumption C_t, while Y_t is the nominal income in period t. W_t and W_{t+1} are nominal stock of financial wealth at period t and t+1. P_t is the price level at period t. β denotes the subjective discount factor over time, and is the corresponding discount rate with the relation $\beta=1/(1+\theta\in[0,1])$. The solution can be obtained by deriving the first order condition, resulting in the "Euler equation":

$$1 = E_t\left[\frac{\beta U'(C_{t+1})}{U'(C_t)}\frac{P_t}{P_{t+1}} \cdot R_{t+1}\right]. \qquad (10)$$

Comparing (10) with (5): $1=E_t[M_{t+1}.R_{t+1}]$, we can see that C-CAPM is actually a special case of SDF with a specific form of stochastic discount factor:

$$M_{t+1} = \frac{\beta U'(C_{t+1})}{U'(C_t)}\frac{P_t}{P_{t+1}}$$

$$= \frac{\beta[U'(C_t)+U''(C_t)\cdot(C_{t+1}-C_t)]}{U'(C_t)}\frac{P_t}{P_{t+1}}$$

$$= \beta\left[1-\left(-\frac{U''(C_t)\cdot C_t}{U'(C_t)}\right)\cdot\left(\frac{C_{t+1}-C_t}{C_t}\right)\right]\frac{P_t}{P_{t+1}}$$

$$= \beta(1-\gamma_t \cdot \frac{\Delta C_{t+1}}{C_t})\frac{P_t}{P_{t+1}} \approx \beta(1-\gamma_t \cdot \Delta \ln C_{t+1})\frac{P_t}{P_{t+1}}$$

(11) (Taylor expansion)

The result follows by combining these two parts:

$\gamma_t = -C_t \cdot U''(C_t)/U'(C_t) > 0$ is the coefficient of relative risk aversion (CRRA);

$$\Delta \ln C_{t+1}=\ln C_{t+1}- \ln C_t=\ln(C_{t+1}/C_t)=\ln(1+ \Delta C_{t+1}/C_t) \approx (C_{t+1}- C_t)/C_t$$

Take logarithm on both sides and use approximation:

$$m_{t+1}=\ln M_{t+1}\approx -\theta- \gamma_t \cdot \Delta \ln C_{t+1}- \Delta \ln P_{t+1}$$

$$=-\theta- \gamma_t \cdot \Delta C_{t+1}- \Delta p_{t+1} \qquad (12)$$

Under the assumption of jointly lognormal distribution of R_{t+1} and M_{t+1} (or equivalently ΔC_{t+1}), we can obtain a further version of no arbitrage condition by substituting (12) in (9):

$$E_t[r_{t+1}-r_t^f]+\frac{1}{2}Var_t[r_{t+1}]=\gamma_t \cdot Cov_t[\Delta c_{t+1},r_{t+1}]+Cov_t[\Delta p_{t+1},r_{t+1}]. \qquad (13)$$

This is a two factor model since there are two explanatory variables, ΔC_{t+1} and Δp_{t+1}. For simplicity, a widely used additional assumption is power utility function which has a constant CRRAγ: $U(C_t) = (C_t^{1-\gamma} -1)/(1-\gamma)$.

(3) CAPM (with monetary model assumption):

In contrast to the dynamic approach in C-CAPM, CAPM puts asset pricing into a *static* optimization problem for a risk averse investor. The investor considers now only one period but many risky assets to tradeoff between risk and return, searching for an optimal choice on the Markowitz efficient frontier to maximize the expected utility function. Different from intertemporal utility maximization, the investment strategy in CAPM is to find an optimal allocation of wealth in different assets rather than an optimal allocation of wealth in different dates.

To obtain the optimal choice, we follow the routine to derive the first order condition. Finally, the solution is found on the *Capital Market Line*:

$$E_t[r_{t+1}^p-r_t^f]=\gamma_t \cdot V_t[r_{t+1}^p]$$

r_{t+1}^p is the rate of return of the optimal portfolio. If there are only risky assets, then it is just the market portfolio r_{t+1}^m defined as the growth rate $(\Delta W_{t+1}/W_t)$ of nominal wealth during period t and t+1. The solution is transformed as follows:

$$E_t[r_{t+1}^m-r_t^f]=\gamma_t \cdot V_t[r_{t+1}^m]$$

The other result of CAPM relates the expected excess return of a risky asset to the expected excess return of the market portfolio. (*Security Market Line*)

$$E_t[r_{t+1}-r_t^f]=\beta_t \cdot E_t[r_{t+1}^m-r_t^f]=\frac{Cov_t[r_{t+1}^m,r_{t+1}]}{V_t[r_{t+1}^m]} \cdot E_t[r_{t+1}^m-r_t^f]$$

β_t is the market beta, which describes the relation between individual asset and the market. The beta for risk-free asset is zero and the beta for the market portfolio is one. Market beta varies over time and across assets: Big diversified companies tend to track the market and have betas close to one, while new technology companies are more risky and have higher betas. Combine these two results together, and then we can write the no arbitrage condition:

$$E_t[r_{t+1}-r_t^f]=\gamma_t \cdot Cov_t[r_{t+1}^m,r_{t+1}]$$

$$=\gamma_t \cdot Cov_t[\frac{\Delta W_{t+1}}{W_t},r_{t+1}] \qquad (14)$$

$$\approx \gamma_t \cdot Cov_t[\Delta \ln W_{t+1},r_{t+1}]$$

r_{t+1}^m can be seen as the market return on nominal wealth, since it

is obtained by the market portfolio. We can assume that the market portfolio consists of hedged and unhedged currency and so the uncertain element in r_t^m is the future spot exchange rate. To push equation (14) further, we can utilize the macroeconomic theory to embody the model in finance.

Monetary model of the exchange rate, which is based on long-term equilibrium, can be employed to provide the observable macroeconomic factors. The exchange rate is determined by future expected relative money supplies and output levels, such as models in Frenkel and Obstfeld and Rogoff. The general idea can be expressed by the plot which you could find in any Macroeconomics textbook[14]:

From this Graph 1, we can see the mechanism of determination of r_{t+1}^m in the money market. The market return, or average interest rate in other words, is negatively related with money supply (MS_{t+1}), but positively related with output level (Y_{t+1}). Hence, r_{t+1}^m can be decomposed into these two macroeconomic factors and approximated as a linear function of money supply and output level. Algebraically, we can write this relation in the logarithmic form:

$$r_{t+1}^m = \kappa \cdot y_{t+1} - \eta \cdot ms_{t+1} + \xi_{t+1} = m_{t+1} \qquad (15)$$

The two coefficients K and are all positive constant. All the lowercases denote the logarithm of the variables. For theoretical convenience, we assume the variables have a linear relation. This condition is a common assumption in most exchange rate determination model. Thus, equation (14) can now be developed to a testable model[15] using equation (15):

$$E_t[r_{t+1} - r_t^f] = \gamma_t \cdot Cov_t[r_{t+1}^m, r_{t+1}]$$
$$= \gamma_t \kappa \cdot Cov_t[y_{t+1}, r_{t+1}] - \gamma_t \eta \cdot Cov_t[ms_{t+1}, r_{t+1}] \qquad (16)$$

Now we can explore the interrelationship between SDF, C-CAPM, and CAPM by comparing equation (9), (13), and (16) It is clear that C-CAPM and CAPM are all special cases of SDF, with different

Graph 1: Mechanism of determination of exchange rate in the money market.
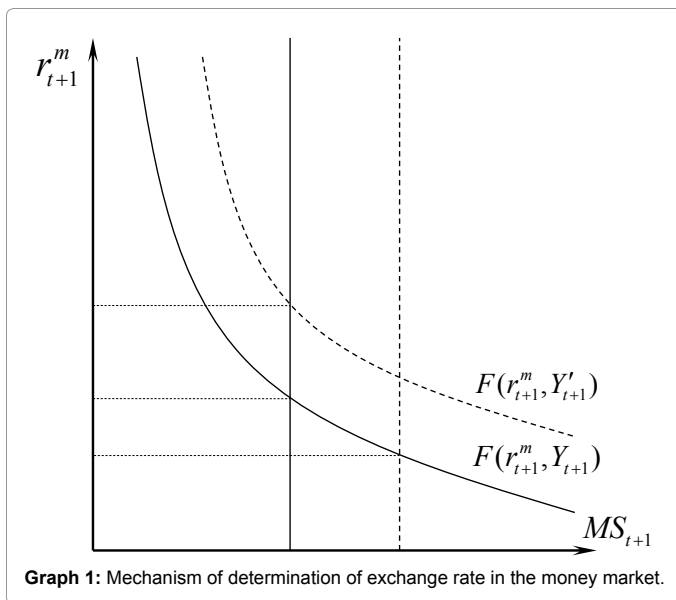
[14]This graph is to be found in the section of derivation of LM schedule. It shows that the output level and the interest rate are positively related, so LM curve has an upwards slope.

[15]Strictly speaking, it should also include a term $\gamma_t \cdot Cov_t[\xi_{t+1}, r_{t+1}]$, but it is supposed to be zero because the error term ξ_{t+1} is uncorrelated with the return of a risky asset.

specifications of stochastic discount factor in the form of conditional covariance. Both models are linear functions of two observable macroeconomic factors, so they are two-factor affine SDF models. Moreover, the difference between C-CAPM and CAPM consists in whether the element is real or nominal. If real consumption is proportional to nominal wealth, then the results of the two models will be identical.

(4) Latent variable affine factor model:

C-CAPM and CAPM share a common feature that the factors of the models are observable. However, there exists another type of affine model with latent factors as its elements. Two most famous affine models with latent variables are Vasicek and Cox-Ingersoll-Ross, assuming that the log of the SDF can be expressed by a linear function of unobservable random variables, which could be substituted by postulated proxy variables in empirical studies. The common assumptions of the two models are lognormality and mean reverting $AR(1)$ process of the discount factors, whereas CIR model distinguishes itself from Vasicek by positive factor constraint. The two latent factor affine models (single factor models) can be compared together because they are generic in structure. As shown below, risk premium in Vasicek model is a constant over time, whereas that in CIR model is time varying. Actually, the latent factors of Vasicek and CIR can both be proxied by future spot interest rate, so CIR model seems more flexible comparatively (Table 4).

Furthermore, these single factor affine models can be extended to multiple factor affine models, comprising more than one latent variable in the linear function of SDF. For example, there might be two latent variables $Z_{1,t+1}$ and $Z_{2,t+1}$, such as the two factor CIR model in Backus et al. [13]. They select *spot interest rates* of domestic and foreign countries as the proxies for the two unobservable variables in their model. The risk premium turns out to be:

$$E_t[r_{t+1} - r_t^f] + \frac{1}{2}V_t[r_{t+1}] = -\frac{1}{2}\lambda^2\sigma^2(z_{1,t+1} - z_{2,t+1})$$

As a summary of pricing theory section, analysis of the underlying relationship among these models is due to be made. The purpose of all the pricing models is to specify the form of the stochastic discount factor, which thus is regarded as the "pricing kernel". They, therefore, are all based on the most general SDF model and taking the form of conditional covariance. So far, all the models concerned are affine model because the expressions of SDF can be written as a linear function of random variables (observable or latent) with error term:

$$M_{t+1} = \beta_0 + \beta_1 \cdot z_{1,t+1} + \cdots + \beta_1 \cdot z_{k,t+1} + \xi_{t+1} = z'_{t+1}\beta + \xi_{t+1}$$

β denotes a column vector of coefficients, sometimes including a constant as the intercept. Z_{t+1} is the vector of explanatory variables, sometimes including unity. ξ_{t+1} denotes the stochastic disturbance with zero mean due to the pure measurement error, so it is supposed to be uncorrelated with Z_{t+1}. Under the assumption of lognormality, the equation is developed as:

Model	Vasicek Model	CIR Model
Assumptions:	$-m_{t+1} = z_{t+1} + \lambda\sigma\varepsilon_{t+1}$ $z_{t+1} = (1-\varphi)\mu + \varphi z_t + \sigma\varepsilon_{t+1}$ $(\varepsilon_{t+1} \sim iid[0,1])$	$-m_{t+1} = z_{t+1} + \lambda\sigma\sqrt{z_t}\varepsilon_{t+1}$ $z_{t+1} = (1-\varphi)\mu + \varphi z_t + \sigma\sqrt{z_t}\varepsilon_{t+1}$ $(\varepsilon_{t+1} \sim iid[0,1])$
Risk premium: $E_t[r_{t+1} - r_t^f] + \frac{1}{2}V_t[r_{t+1}]$	$-\frac{1}{2}\lambda^2\sigma^2$	$-\frac{1}{2}\lambda^2\sigma^2 z_{t+1}$

Table 4: Two latent factor affine models.

$$m_{t+1} = \ln M_{t+1} = z'_{t+1}\beta + \xi_{t+1}$$

Where Z_{t+1} could be $[1, \Delta c_{t+1}, \Delta p_{t+1}]'$ as in C-CAPM, $[y_{t+1}, ms_{t+1}]'$ as in CAPM, or r_{t+1} as in Vasicek and CIR latent factor models. The different specifications for SDF then can be uniformed as follows:

C-CAPM: $m_{t+1} = -\theta - \gamma_t \cdot \Delta c_{t+1} - \Delta p_{t+1} + \xi_{t+1} = z'_{t+1}\beta + \xi_{t+1}$

($\beta = [-\theta, -\gamma_t, -1]', z_{t+1} = [1, \Delta c_{t+1}, \Delta p_{t+1}]'$, and $E_t[\xi_{t+1}|z_{t+1}] = 0$)

CAPM: $m_{t+1} = \gamma_t \cdot r^m_{t+1} + \xi_{t+1} = \gamma_t \kappa \cdot y_{t+1} - \gamma_t \eta \cdot ms_{t+1} + \xi_{t+1} = z'_{t+1}\beta + \xi_{t+1}$

($\beta = [\beta_1, \beta_2]', z_{t+1} = [\Delta y_{t+1}, \Delta ms_{t+1}]'$, and $E_t[\xi_{t+1}|z_{t+1}] = 0$)

Vasicek: $m_{t+1} = -z_{t+1} - \lambda \sigma \varepsilon_{t+1} = -r_{t+1} - \lambda \sigma \varepsilon_{t+1} = \beta \cdot z_{t+1} + \xi_{t+1}$

($\beta = -1, z_{t+1} = r_{t+1}$, and $\xi_{t+1} = -\lambda \sigma \varepsilon_{t+1} \sim iid[0, \lambda^2 \sigma^2]$)

CIR: $m_{t+1} = -z_{t+1} - \lambda \sigma \sqrt{z_t} \varepsilon_{t+1} = -r_{t+1} - \lambda \sigma \sqrt{z_t} \varepsilon_{t+1} = \beta \cdot r_{t+1} + \xi_{t+1}$

($\beta = -1, z_{t+1} = r_{t+1}$, and $\xi_{t+1} = -\lambda \sigma \sqrt{z_t} \varepsilon_{t+1} \sim iid[0, \lambda^2 \sigma^2 z_t]$)

In the light of this argument, all the SDF models can be expressed in the form of linear function, which is the basis of econometric modeling.

Market structure: Logically, there are three types of market structures [12] in terms of influence of domestic and foreign investors on exchange rate. It is both necessary and important to find out the underlying relationship, which is another building block of asset pricing models.

(1) Domestic investor model: Only domestic investors affect the exchange rate as a result of their purchase of the foreign asset. No arbitrage condition can be derived from risk adjusted UIP solely in domestic market:

$$\text{excess return} = \Re_{t+1} = (\Delta s_{t+1} + r^*_t) - r_t \qquad (17)$$

Here $(\Delta s_{t+1} + r^*_t)$ denotes the foreign risk free rate in terms of domestic currency, and r_t is the domestic risk free rate. The difference between them is just the excess return. Combine (17) and CIP condition: $f_{t,t+1} = s_t + r_t - r^*_t$.

$$\begin{cases} E_t[\Re_{t+1}] = (E_t[\Delta s_{t+1}] + r^*_t) - r_t = E_t[s_{t+1} - f_{t,t+1}] \\ Var_t[\Re_{t+1}] = Var_t[\Delta s_{t+1}] \end{cases} \qquad (18)$$

Under the assumption of log normality, we have obtained in equation (9):

$$E_t[r_{t+1} - r^f_t] + \frac{1}{2}Var_t[r_{t+1}] = -Cov_t[m_{t+1}, r_{t+1}]$$

Now the excess return we can get the no arbitrage condition in the FOREX market under domestic investor model by combining equation (18) and (9):

$$E_t[\Re_{t+1}] = -\frac{1}{2}Var_t[\Delta s_{t+1}] - Cov_t[m_{t+1}, \Delta s_{t+1}] \cdot \qquad (19)$$

(2) Foreign investor model: Only foreign investors affect the exchange rate as a result of their purchase of domestic asset. Similar to domestic investor model, we can obtain the no arbitrage condition[16] following the former routine:

$$E_t[\Re_{t+1}] = \frac{1}{2}Var_t[\Delta s_{t+1}] - Cov_t[m^*_{t+1}, \Delta s_{t+1}] \qquad (20)$$

(3) Domestic and foreign investor model: In practice, both domestic investors and foreign investors will carry out the FOREX trade, so both sides will affect the market. To get the no arbitrage condition under this

---

[16]We write the condition in the form of US investor risk premium, using the relations:

$\Delta s_{t+1} = -\Delta s^*_{t+1}$ and $\Re_{t+1} = -\Re^*_{t+1}$.

---

assumption, domestic and foreign investor equations are used to get a combined model:

$$E_t[\Re_{t+1}] = -Cov_t\left[\frac{m^*_{t+1} + m_{t+1}}{2}, \Delta s_{t+1}\right] \cdot \qquad (21)$$

Simultaneously, we can get another relation by subtracting (19) from (20):

$$Var_t[\Delta s_{t+1}] = Cov_t[m_{t+1} - m^*_{t+1}, \Delta s_{t+1}]$$

Variance of a random variable is equal to covariance between this variable and another random variable if and only if the two variables are the same in mean. Hence, we can get an important relation from (22):

$$m^*_{t+1} = m_{t+1} + \Delta s_{t+1} . \qquad (22)$$

Equation (23) implies that three no arbitrage conditions based on domestic, foreign, and both investors are identical because there is a linear relation between domestic and foreign SDF. Hence, we could then have only one equation from (19), (20), and (21) to estimate; the result could be expressed equivalently in any of the three ways through the Relation (23).

Empirical model: Suppose that the domestic investor is US based, and the foreign investor is UK based. The exchange rate is the sterling dollar exchange rate, which is the number of dollars per pound. Suppose the joint distributions of the future spot rate and other relevant random variables are lognormal. To push our analysis further, we have to combine the two former parts together to get testable econometric models and compare them for our model selection purpose.

There are two popular econometric tools to model the risk premium in FOREX market, i.e. Vector Autoregressive (VAR), and Multivariate GARCH in Mean (MGM). VAR restricts a coefficient matrix to satisfy the no arbitrage condition, when a vector of returns is used. However, it is not a valid way because the specifications of risk premia involve conditional covariance terms that VAR does not include. In contrast, Multivariate GARCH in Mean (MGM) is more general. Smith et al. suggest that the conditional covariance structure of FOREX market can be well approximated by ARCH process. Thus, the joint conditional distribution of exchange rate can be modeled by a MGM process. This postulation allows the conditional mean of the distribution to be affected by lagged values and by the conditional covariance matrix.

$$\begin{cases} \text{VAR}: v_{t+1} = \alpha + \Gamma v_t + \varepsilon_{t+1} \\ \text{MGM}: v_{t+1} = \alpha + \Gamma v_t + \Phi g_t + \varepsilon_{t+1} \end{cases}$$

$$v_{t+1} = \{\Re_{t+1}, z'_{t+1}\}' = \{\Re_{t+1}, z'_{t+1}\}' .$$

$\varepsilon_{t+1}|I_t \sim N[0, H_{t+1}]$, where I_t is the information set available at period t.

$g_t = vech[H_{t+1}]$ where $vech[.]$ is an operator converting the lower triangle of a symmetric matrix into a vector.

Now we can make use of MGM to combine the two building blocks, i.e. asset pricing models and FOREX market structure. The dependent variable is the same across models (i.e. excess return $\Re_{t+1}$), and the error term ε_{t+1} follows a GARCH (1,1) process. The models only differ in the vector of regressors.

【C-CAPM】

US investor model:

$$\Re_{t+1} = -\frac{1}{2}Var_t[\Delta s_{t+1}] + \gamma_t Cov_t[\Delta s_{t+1}, \Delta c_{t+1}^{US}] + Cov_t[\Delta s_{t+1}, \Delta p_{t+1}^{US}] + \varepsilon_{t+1}$$

UK investor model:

$$\Re_{t+1} = \frac{1}{2}Var_t[\Delta s_{t+1}] + \gamma_t^* Cov_t[\Delta s_{t+1}, \Delta c_{t+1}^{UK}] + Cov_t[\Delta s_{t+1}, \Delta p_{t+1}^{UK}] + \varepsilon_{t+1}$$

US and UK investor model:

$$\Re_{t+1} = \frac{1}{2}\gamma_t Cov_t[\Delta s_{t+1}, \Delta c_{t+1}^{US}] + \frac{1}{2}Cov_t[\Delta s_{t+1}, \Delta p_{t+1}^{US}]$$
$$+ \frac{1}{2}\gamma_t^* Cov_t[\Delta s_{t+1}, \Delta c_{t+1}^{UK}] + \frac{1}{2}Cov_t[\Delta s_{t+1}, \Delta p_{t+1}^{UK}] + \varepsilon_{t+1}$$

General alternative model:

$$\Re_{t+1} = constant + \beta_1 \Re_t + \beta_2(s_{t+1} - f_{t,t+1}) + \beta_3 Var_t[\Delta s_{t+1}]$$
$$+ \beta_4 Cov_t[\Delta s_{t+1}, \Delta c_{t+1}^{US}] + \beta_5 Cov_t[\Delta s_{t+1}, \Delta p_{t+1}^{US}]$$
$$+ \beta_6 Cov_t[\Delta s_{t+1}, \Delta c_{t+1}^{UK}] + \beta_7 Cov_t[\Delta s_{t+1}, \Delta p_{t+1}^{UK}] + \varepsilon_{t+1}$$

【CAPM】

US investor model:

$$\Re_{t+1} = -\frac{1}{2}Var_t[\Delta s_{t+1}] - \gamma_t \kappa Cov_t[\Delta s_{t+1}, \Delta y_{t+1}^{US}] + \gamma_t \eta Cov_t[\Delta s_{t+1}, \Delta ms_{t+1}^{US}] + \varepsilon_{t+1}$$

UK investor model:

$$\Re_{t+1} = \frac{1}{2}Var_t[\Delta s_{t+1}] - \gamma_t^* \kappa^* Cov_t[\Delta s_{t+1}, \Delta y_{t+1}^{UK}] + \gamma_t^* \eta^* Cov_t[\Delta s_{t+1}, \Delta ms_{t+1}^{UK}] + \varepsilon_{t+1}$$

US and UK investor model:

$$\Re_{t+1} = -\frac{1}{2}\gamma_t \kappa Cov_t[\Delta s_{t+1}, \Delta y_{t+1}^{US}] + \frac{1}{2}\gamma_t \eta Cov_t[\Delta s_{t+1}, \Delta ms_{t+1}^{US}]$$
$$- \frac{1}{2}\gamma_t^* \kappa^* Cov_t[\Delta s_{t+1}, \Delta y_{t+1}^{UK}] + \frac{1}{2}\gamma_t^* \eta^* Cov_t[\Delta s_{t+1}, \Delta ms_{t+1}^{UK}] + \varepsilon_{t+1}$$

General alternative model:

$$\Re_{t+1} = constant + \beta_1 \Re_t + \beta_2(s_{t+1} - f_{t,t+1}) + \beta_3 Var_t[\Delta s_{t+1}]$$
$$+ \beta_4 Cov_t[\Delta s_{t+1}, \Delta y_{t+1}^{US}] + \beta_5 Cov_t[\Delta s_{t+1}, \Delta ms_{t+1}^{US}]$$
$$+ \beta_6 Cov_t[\Delta s_{t+1}, \Delta y_{t+1}^{UK}] + \beta_7 Cov_t[\Delta s_{t+1}, \Delta ms_{t+1}^{UK}] + \varepsilon_{t+1}$$

【Vasicek model】

US investor model: $\Re_{t+1} = -\frac{1}{2}\lambda^2\sigma^2 - \frac{1}{2}Var_t[\Delta s_{t+1}] + \varepsilon_{t+1}$

UK investor model: $\Re_{t+1} = \frac{1}{2}\lambda^{*2}\sigma^{*2} - \frac{1}{2}Var_t[\Delta s_{t+1}] + \varepsilon_{t+1}$

US and UK investor model: $\Re_{t+1} = -\frac{1}{2}\lambda^2\sigma^2 + \frac{1}{2}\lambda^{*2}\sigma^{*2} - Var_t[\Delta s_{t+1}] + \varepsilon_{t+1}$

General alternative model:

$$\Re_{t+1} = constant + \beta_1 \Re_t + \beta_2(s_{t+1} - f_{t,t+1}) + \beta_3 Var_t[\Delta s_{t+1}] + \varepsilon_{t+1}$$

【CIR model】

US investor model: $\Re_{t+1} = -\frac{1}{2}Var_t[\Delta s_{t+1}] - \frac{1}{2}\lambda^2\sigma^2 r_t^{US} + \varepsilon_{t+1}$

UK investor model: $\Re_{t+1} = -\frac{1}{2}Var_t[\Delta s_{t+1}] + \frac{1}{2}\lambda^{*2}\sigma^{*2} r_t^{UK} + \varepsilon_{t+1}$

US and UK investor model: $\Re_{t+1} = -Var_t[\Delta s_{t+1}] - \frac{1}{2}\lambda^2\sigma^2 r_t^{US} + \frac{1}{2}\lambda^{*2}\sigma^{*2} r_t^{UK} + \varepsilon_{t+1}$

General alternative model:

$$\Re_{t+1} = constant + \beta_1 \Re_t + \beta_2(s_{t+1} - f_{t,t+1}) + \beta_3 Var_t[\Delta s_{t+1}] + \beta_4 r_t^{US} + \beta_5 r_t^{UK} + \varepsilon_{t+1}$$

Evidence: Monthly data of industrial production is used for output level, retail sales for consumption, CPI for price level, and M0 for money supply. The spot interest rates for US and UK are annualized. We use these substitutions because the macroeconomic data with a frequency higher than quarterly is scarce.

The data spans 20 years from 1st September 1986 to 1st September 2016, which are the latest resources from DATASTREAM.

Due to the presence of conditional variance and covariance in the equations, they are no longer the linear regression model. We shall have to bother some more complicated econometric method such as General Method of Moments (*GMM*), which is a nonlinear instrumental variables estimation method. It is widely employed in estimating affine and nonlinear models. The use of GMM in financial time series data is supported by its robustness to heteroskedasticity. The principle is that there exists a set of true parameters for which the vector of error terms is orthogonal to the set of instrumentals. Based on this condition, we can construct a moment equation and a corresponding empirical moment equation, the elements of which is just the counterparts of the moment equation in empirical world. The consistent and efficient GMM estimator is obtained by minimizing the objective empirical moment equation in two steps. Meanwhile, to check the specification for model estimated by GMM, J-test of over-identification restrictions is proposed by Hansen. We will utilize the two complementary econometric tools to carry out our tests.

The selected results of the empirical models are listed in Tables 5 and 6; here the results for C-CAPM and CAPM are reported. The rows are the coefficients of regressors which might appear in different models, and the columns are different models based on the FOREX market structure as well as the general alternative model. As suggested by Smith and Wickens, the coefficients of $Var_t[\Delta s_{t+1}]$ for US and UK models are imposed instead of being estimated. It is not 1/2 as postulated in the original models because the data are all annualized[17]. The general alternative model comprises not only all the variables in the three models, but also some lagged variables.

(1) Latent variable affine factor models: Compared to other SDF models, we can easily reject them by theoretical comparison and brief literature review, so I did not carry out the tests for latent variable affine factor models. Most of the empirical evidence strongly suggests that the risk premium is time varying. Vasicek model, however, implies a constant risk premium over time, i.e. $-\lambda^2\sigma^2/2$. CIR model has similar problem, despite its superficial adjustment for a changeable risk premium with the latent variable, i.e. $-\lambda^2\sigma^2 z_{t+1}/2$. Based on this argument, the failure of latent factor affine models is predictable. As shown earlier in Dai and Singleton and Backus et al, even when we extend the model to two latent factors, the GMM estimates of the coefficients still display extreme distributional properties for both models. An extended version of the approach is created by Hollifield and Yaron [14]. They adjust the model to allow for the risk premium to be decomposed into real and nominal components for each country and their interaction. The result of this extension also reports failure of the latent variable affine factor models.

(2) C-CAPM: The result for C-CAPM is reported in Table 5. As shown in the two columns of US investor model and UK investor model, the conditional covariance terms are significant at 10% level. In the combined model, only the conditional covariance between the excess return and US consumption growth rate is significant. In the general model, however, none of the conditional covariance terms is significant.

---

[17] Since the data are monthly, when they are annualized, the coefficient are changed as well:

$$\frac{1}{2} \xrightarrow{annualization} \frac{1}{2} \times \frac{1}{12}\% = 0.0004166 \approx 0.0004$$

Regressor	US	UK	US-UK	General
Constant	-	-	-	-0.287 (0.029)
$\Re_t$	-	-	-	**0.351** **(5.24)**
$(s_{t+1} - f_{t,t+1})$	-	-	-	**1.689** **(2.81)**
$Var_t[\Delta s_{t+1}]$	*-0.0004*	*0.0004*	-	-0.003 (0.37)
$Cov_t[\Delta s_{t+1}, \Delta c_{t+1}^{US}]$	**0.027** **(2.13)**	-	**0.024** **(2.10)**	0.022 (1.41)
$Cov_t[\Delta s_{t+1}, \Delta p_{t+1}^{US}]$	-0.097 (1.64)	-	-0.032 (0.23)	-0.063 (0.35)
$Cov_t[\Delta s_{t+1}, \Delta c_{t+1}^{UK}]$	-	-0.015 (1.91)	-0.013 (1.19)	-0.014 (1.45)
$Cov_t[\Delta s_{t+1}, \Delta p_{t+1}^{UK}]$	-	**0.040** **(2.04)**	-0.015 (0.34)	0.005 (0.072)
Conditional Variance of the regression				
Constant	**1321.7** **(3.85)**	**1295.4** **(3.91)**	**1304.9** **(2.98)**	**1024.5** **(4.97)**
ε_t^2	**0.642** **(5.44)**	**0.599** **(5.25)**	**0.642** **(5.48)**	**0.668** **(5.73)**
ε_t^2	**0.273** **(3.03)**	**0.341** **(3.15)**	**0.329** **(2.92)**	**0.179** **(2.91)**

Table 5: Estimates of the C-CAPM for FOREX market efficiency (1986-2016). This nonlinear regression is based on MGM, using GMM estimation. The dependent variable is excess return $\Re_{t+1}$. The test for GARCH process of the conditional variance σ_{t+1}^2 is carried out as well. Numbers in parentheses are t-statistics. The figures in bold font are significant and the others are insignificant at 10% level.

Moreover, after the adjustment[18], all of the estimated values of CRRA and other coefficients have wrong signs or are very large, compared to our postulated values. For example, the implied estimated coefficients of CRRA for US investor is -276, -282, and -391 in the US investor based, UK investor based, and combined models respectively. The signs for the coefficient of covariance of excess return with inflation are correct in the three models, but the size of the coefficients is far too large.

The empirical foundings here are similar to those of Smith and Wickens, Mark and Wu, Engel, and Lewis. In conclusion, general equilibrium C-CAPM model is rejected by the latest data and the forward premium puzzle is not resolved. Hence, the theory does not seem to be able to provide a satisfactory basis for our practical purpose either.

(3) CAPM: The model used here actually incorporates the monetary model in Macroeconomics into the traditional CAPM. The results are reported in Table 6 with a similar form of C-CAPM. The theoretical predictions for the monetary model are that the coefficient on conditional covariances should be positive for US money and negative for US output, and these sings should be reversed for the UK variables. The estimates for the US investor model and the UK investor model have the correct signs and are significant. For the combined model, all the signs are correct, but the covariances of excess return with UK money supply are not significant. In the general model, the covariances with output growth and the lagged variables are comparatively significant.

These results therefore provide considerable support for the monetary CAPM despite that the "forward premium puzzle" still remains due to the significance of lagged excess return and forward

premium in the general model. Anyway, monetary CAPM is the "best" models in terms of our practice oriented criterion among all those models we have gone through. Furthermore, the UK investor model is outstanding among other CAPM models due to its high significance.

To conclude for the model selection section, we are happy to find out the best fitting model according to the latest data. As a short-term partial equilibrium model, monetary Capital Asset Pricing Model incorporates macroeconomic variables into asset pricing which receives the most support in empirical tests. Certainly, it is not the end of our research because this result is just based on the information available and the specific practical criterion. New data and theoretical development might advance our "best" model in the future.

Application

To realize the value of the results obtained in the model selection, we consider how to use CAPM in investment. The first thing to start with is to outline the FOREX market in a practical sense. A new approach based on our earlier discussion is proposed at the end of the paper. Comparison of the two approaches is made by an experiment. It turns out the new approach considerably reduces the risk related to FOREX market volatility.

FOREX market in practice

The FOREX market exists wherever one currency is traded for another. It is by far the largest market in the world in terms of cash value traded. The transaction happening in FOREX markets across the world currently exceeds $1.9 trillion per day (on average). These huge trading volume and high liquidity of FOREX assets guarantees no arbitrage opportunity and efficiency of the market.

The transactions happened in FOREX market are incurred by all kinds of participants include governments, central banks, banks, currency speculators, multinational corporations, and other financial institutions such as investment management firms. They typically

---

[18]The estimated values of coefficients need to be adjusted to be in line with the corresponding counterparts in our models. It is because the estimates in Table 1 are for the coefficients of the product of the conditional standard deviations, not for the conditional covariance itself as in the theory.

Regressor	US	UK	US-UK	General
Constant	-	-	-	17.458 (1.10)
$\Re_t$	-	-	-	**0.342** **(5.34)**
$(s_{t+1}-f_{t,t+1})$	-	-	-	**1.728** **(2.32)**
$Var_t[\Delta s_{t+1}]$	*-0.0004*	*0.0004*	-	-0.014 (0.05)
$Cov_t[\Delta s_{t+1}, \Delta ms_{t+1}^{US}]$	**0.156** **(1.69)**	-	**0.132** **(1.95)**	-0.274 (0.71)
$Cov_t[\Delta s_{t+1}, \Delta y_{t+1}^{US}]$	-0.051 (1.62)	-	**-0.132** **(2.93)**	**-0.092** **(1.93)**
$Cov_t[\Delta s_{t+1}, \Delta ms_{t+1}^{UK}]$	-	**-0.091** **(2.25)**	-0.043 (0.75)	-0.031 (0.84)
$Cov_t[\Delta s_{t+1}, \Delta y_{t+1}^{UK}]$	-	**0.044** **(2.63)**	**0.058** **(4.19)**	**0.042** **(3.01)**
Conditional Variance of the regression				
Constant	**1208.3** **(2.79)**	**1287.5** **(2.97)**	**986.6** **(4.37)**	**897.2** **(4.34)**
σ_t^2	**0.652** **(4.31)**	**0.651** **(4.64)**	**0.699** **(3.97)**	**0.774** **(6.91)**
ε_t^2	**0.254** **(2.28)**	**0.239** **(2.74)**	**0.247** **(2.87)**	**0.168** **(2.58)**

Table 6: Estimates of the CAPM for FOREX market efficiency (1986-2016). The empirical model combines the traditional CAPM with the monetary model, consisting of observable macroeconomic variables such as output level and consumption level. It is also based on MGM model, using GMM estimation. The dependent variable is excess return $\Re_{t+1}$. The test for GARCH process of the conditional variance σ_{t+1}^2 is carried out as well. Numbers in parentheses are t-statistics. The figures in bold font are significant and the others are insignificant at 10% level.

manage large accounts of fund on behalf of customers such as pension funds and endowments. FOREX market is used to facilitate transactions in foreign securities. An investment manager with an international equity portfolio will need to buy and sell foreign currencies in the spot market in order to pay for purchases of foreign equities. Since the FOREX transactions are secondary to the actual investment decision, they are not seen as speculative or aimed at profit maximization. Thus, financial derivatives such as forward rates, future and swap contracts are usually utilized to reduce uncertainty incurred by the volatile exchange rates.

Some investment management firms also have more speculative specialist currency overlay operations, which manage clients' currency exposures with the aim of generating profits as well as limiting risk. Whilst the number of this type of specialist firms is quite small, many have a large value of assets under management, and hence can generate large trades. They regard the currency itself as a risky asset, which is all the same as other assets. Hence, ordinary asset allocation method can be directly used. CAPM, as we have just proved, is the "best" benchmark model in FOREX market up to now. However, our focus is on the former type of international asset allocation.

CAPM analysis in practice: Although many studies tend to reject CAPM in theoretical settings, it is still the most popular model for fund management in financial markets. To obtain the optimal portfolio, a risk averse investor should tradeoff between expected return and return volatility. The optimal portfolio lies on the Capital Market Line (CML) which we bothered earlier.

Traditional methods of portfolio selection such as the mean-variance analysis of Markowitz and the CAPM due to Sharpe and Lintner are based on the assumption of constant asset return volatility and thus a constant portfolio frontier. In practice, however, it is noted that the covariance matrix of returns, from which the portfolio frontier is formed, is actually time-varying [15]. Ferson and Harvey exploit

this in their analysis of asset pricing. This implies that instead of the portfolio frontier being based on the unconditional covariance matrix of returns, which is then constant, it should be calculated from the conditional covariance matrix, which is changing over time [16-18].

New methodology: Based on the time varying CAPM, Flavin and Wickens suggest that investors in UK assets could enjoy a significant reduction in portfolio risk by using a time varying conditional covariance matrix. As the frontier is also time varying, the investor needs to continuously rebalance the portfolio. Moreover, multivariate GARCH (1,1) model is once again employed to simulate the volatility. It is also suggested that macroeconomic variables can be incorporated into traditional CAPM to account for both the conditional mean and the conditional covariance matrix of the asset returns [19-21].

To take account of the effect of macroeconomic variables on portfolio selection, we have to incorporate the macroeconomic variables into the vector of excess returns, as we have achieved in model selection. Compared to our earlier modeling in (16), we replace output and money supply by single factor inflation. It is actually an equivalent transformation since the three variables have a linear relation. The conditional covariance matrix of excess returns of this joint distribution is formed from the multivariate marginal conditional distribution of the excess returns. Hence, the conditional distribution of excess returns will depend on the volatility of the macroeconomic variables, which can be used to help predict the covariance matrix of the excess returns and then the portfolio frontier. The construction of the optimal portfolio is now obtained as before, but the asset shares will differ from those computed without taking account of the macroeconomic variables. That is to say, under the same expected return, the risk of the portfolio should be reduced by our new approach.

Case study: To compare the traditional CAPM and modified CAPM approach, Flavin and Wickens carry out a case study. Recall our earlier model selection, UK investor model is the best fitting model in

practice. Therefore we suppose a US investor is considering allocating her money (USD) on three risky UK assets (UK equity, a long-term UK government bond, and a short-term UK government bond) as well as a risk free asset (30-day Treasury bill). Equity is represented by the FT "All Share Index", long-term government bonds are represented by the FT "British government stock with over 15 years to maturity index", and short-term government bonds are represented by the FT "British government stock with less than 5 years to maturity index". Each is expressed as an excess return over the risk free rate of 30-day Treasury bill.

The portfolio weights for the three UK risky assets calculated by the traditional CAPM method are 70%, 20%, and 10% respectively for equity, long-term bond, and short-term bond. Whereas under the new approach after taking inflation into account, the shares are 74%, 14%, and 12%, significantly changed. The main difference lies in the riskiness of the portfolios. The risk associated with the new portfolio is much lower - on average by almost 24% in each period. We can also examine the portfolio performance to judge the effect of the new approach. A supportive result is evidenced by an increase of 1.6% in the Sharpe Performance Index.

Conclusion

Starting with a simple idea to select a suitable model for practice in FOREX market, this paper examines various prevailing models using the latest data and the advanced econometric tools. According to the practice oriented criterion, long-term model and risk neutral model are rejected, ending up with PPP puzzle and forward premium puzzle respectively. Among the risk adjusted models, I construct a series of empirical SDF models based on multivariate GARCH in mean process and FOREX market structures. Monetary CAPM, which incorporates observable macroeconomic factors into the model, turns out to be the "best" one, especially the UK investor based model.

In the light of this result of model selection, a further application of CAPM is made to the practice. Different from traditional static CAPM, the covariance matrix of return is assumed to be time varying. Meanwhile, macroeconomic variables, such as inflation, are used to account for the mean and volatility of asset return, resulting in a new approach. Compared with the traditional CAPM, volatility of the new portfolio is considerably reduced and the performance of the portfolio is also promoted.

Theory is created to be applied into practice, though it is not the only purpose of theory. This paper is striving to fill the gap between the conceptual world and the real world. That is why it lives up to a practice oriented criterion to select model, which is still a lacuna in existing literature to my knowledge.

Lastly, it is necessary to re-stress that, in this paper, the model selection is pragmatic or practice oriented. The resulting model might not be the best one in theory, and the eliminated models are not worthless as well. Actually some sophisticated models could be sounder in the theoretical sense. However, the main purpose of this paper is not intended to create a delicate theory but to advance practice.

References

1. Chen X, Minford P, Tian K, Zhou P (2016) Who provides the Capital for Chinese Growth: the Public or the Private Sector? Applied Economics.

2. Dai L, Minford P, Zhou P (2015) A DSGE Model of China. Applied Economics 47: 6438-6460.

3. Obstfeld M, Rogoff K (2000) The Six Major Puzzles in International Macroeconomics: Is There a Common Cause? MIT press, USA.

4. Lothian, James R, Taylor Mark P (1996) Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries. Journal of Political Economy 104: 488-509.

5. Frenkel JA (1976) A Monetary Approach to the Exchange Rate: Doctrinal Aspects and Empirical Evidence. Scandinavian Journal of Economics 78: 169-191.

6. Frenkel JA (1986) International Capital Mobility and Crowding-out in the US Economy: Imperfect Integration of Financial Markets or of Goods Markets? How open is the U.S. Economy? pp: 33-67.

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

8. Samuelson P (1964) Theoretical Notes on Trade Problems. Review of Economics and Statistics 46: 146-154.

9. Fama E (1984) Forward and Spot Exchange Rates. Journal of Monetary Economics 14: 319-318.

10. Froot KA, Thaler RH (1990) Anomalies: Foreign Exchange. Journal of Economic Perspectives 4: 179-92.

11. Smith PN, Wickens MR (2002) Asset Pricing with Observable Stochastic Discount Factors. Journal of Economic Surveys 16: 397-446.

12. Smith PN, Sorensen S, Wickens MR (2001) Macroeconomic Sources of FOREX Risk. Mimeo, The University of York, UK.

13. Backus D, Foresi S, Telmer C (2001) Affine Term Structure Models and the Forward Premium Anomaly. Journal of Finance 56: 279-304.

14. Hollifield B, Yaron A (2001) The Foreign Exchange Risk Premium: Real and Nominal Factors. Mimeo, Carnegie Mellon University.

15. Harvey CR (1991) The World Price of Covariance Risk. Journal of Finance 46: 111-157.

16. Balassa B (1964) The Purchasing Power Parity Doctrine: A Reappraisal. Journal of Political Economy 72: 584-596.

17. Burnside C, Eichenbaum M, Rebelo S (2007) The Returns to Currency Speculation. NBER Working Papers.

18. Frenkel JA (1981) The Collapse of Purchasing Power Parity During the 1970s. European Economic Review 16: 145-65.

19. Lewis KK (1994) Puzzles in International Financial Markets. NBER Working Paper No: 4951.

20. Minford P, Xu Y, Zhou P (2015) How Good are Out of Sample Forecasting Tests on DSGE Models? Italian Economic Journal 1: 333-351.

21. Taylor A, Taylor M (2004) The Purchasing Power Parity Debate. NBER Working Paper No 10607.