# Escola dePós-Graduação em Economia -- EPGE <br> Fundação Getúlio Vargas 

# Essays on the Relashionship between the Equity and theForward Premium Puzzles 

Tese submetida à Escola dePós-Graduação emEconomia da Fundação Getúlio Vargas como requisito de obtenção do t́́tulo de Doutor em Economia.

Aluno: Paulo Rogério Faustino Matos Professor Orientador: Carlos Eugênio Ellery Lustosa da Costa

## Livros Grátis

http://www.livrosgratis.com.br
Milhares de livros grátis para download.

# Escola dePós-Graduação em Economia -- EPGE <br> Fundação Getúlio Vargas 

## Essays on the Relashionship between the Equity and theForward Premium Puzzles

Tese submetida à Escol a de Pós-Graduação em Economia da Fundação Getúlio Vargas como requisito de obtenção do título de Doutor em Economia.

Aluno: Paulo Rogério Faustino Matos Banca Examinadora: Carlos Eugênio Ellery Lustosa da Costa (Orientador, EPGE)<br>Caio Ibsen deAlmeida (EPGE)<br>Marco Antônio Bonomo (EPGE)<br>Marcelo Fernandes (Queen Mary University of London)<br>Ricardo Dias de Oliveira Brito (Ibmec-SP)

Contents

# Acknowledgements iv 

List of Tables v

List of Figures vii

Introduction 1

Chapter 1 - The Forward- and the Equity-Premium Puzzles: Two Symptoms of the SamellIness? 3

Chapter 2 - Do Equity and the Foreign Currency Risk Premiums Display Common Patterns? 38

Chapter 3 - Modeling foreign currency risk premiums with timevarying covariances 66

Acknowledgements

Agradeço,
A Deus pel a iluminação e proteção,
À minha família pelo apoio incondicional e confiança,
À minha noiva e futura esposa, Cristiana, pelo amor e compreensão ao longo deste período emque a privei de meu convívio,

Ao meu orientador, Carlos Eugênio, um incansável parceiro cuja contribuição foi fundamental para a minha formação e para o desenvolvimento desta tesee
Ao corpo docente, discente e demais funcionários da EPGE, instituição a qual serei eternamente grato.

## List of Tables

Chapter 1
Table 1 - Data summary statistics 24
Table 2 - Forward-Premium tests for the Canadian government bonds (In sample exercise) 26
Table 3 - Forward-Premium tests for the German govemment bonds (In sample exercise) 27

Table 4 - Forward-Premium tests for the J apanese government bonds (In sample exercise) 28

Table 5 - Forward-Premium tests for the Swiss govemment bonds (In sample exercise) 29
Table 6 - Forward-Premium tests for the Bristish government bonds (In sample exercise) 30

Table 7 - Equity-Premiumtests (In sample exercise) 31
Table 8 - Forward-Premium tests for the Canadian govemment bonds (Ou of sample exercise) 32
Table 9 - Forward-Premium tests for the German government bonds (Out of sample exercise) 33
Table 10 - Forward-Premium tests for the Japanese government bonds (Out of sample exercise) 34

Table 11 - Forward-Premium tests for the Swiss govemment bonds (Out of sample exercise) 35

Table 12 - Forward-Premium tests for the British government bonds (In sample exercise) 36
Table 13 - Equity-Premiumtests (In sample exercise) 37

Chapter 2
Table 1 - Data summary statistics 59
Table 2 - Testing CRRA preference 60
Table 3 - Testing Abl (1990) catching up with theJ oneses preference 61
Table 4 - Testing Abel (1999) keeping up with theJ oneses preference 62
Table 5 - Testing Epstein and Zin $(1989,1991)$ preference 63

Table 6 - Testing Campbell and Cochrane (1999) habit formation preference 64
Table 7 - Testing a generalized Campbell and Cochrane (1999) habit formation preference 65

## Chapter 3

Table 1 - Data summary statistics 83
Table2 - Models of conditional variance 85
Table 3 - Conditional and adjusted regressions 87

List of Figures


## Introduction

The Consumption Capital Asset Pricing Model-CCAPM—presented to the profession by the seminal works of Lucas (1978) and Breeden (1979) specifies from an equilibrium relation the Stochastic Discount Factor-SDF-as the intertemporal marginal rate of substitution for the representative agent, incorporating definitely the financial theory to the theoretical framework developed by the economists of choice under uncertainty.

In its canonical version, the CCAPM defines the consumption growth as the SDF. Although this specification has been extensively used in finance literature due to its convenience, it really does not work well in practice, there being many evidences of its incapability to account for stylized facts. The equity premium-EPPand the forward premium puzzles-FPP-are two of the most famous and reported empirical failures of this consumption-based approach.

The EPP, commonly associated with the works of Hansen and Singleton (1983) and Mehra and Prescott (1985), is how one calls the incapacity of consumption based asset pricing models, with reasonable parameters for the representative agent's preferences, to explain the excess return of stock market with respect to the risk free bond in the United States. The departure point of the puzzle is an attempt to fit the Euler equation of a representative agent for the American economy, which has proven to be an elusive task.

The FPP, on the other hand, relates to the difference between the forward rate and the expected future value of the spot exchange rate in a world with rational expectations and risk neutrality. Once more, this risk premium model fails when it is not able to generate the conditional bias of the forward rates as predictors of future spot exchange rates that characterizes the FPP.

Relaxing the risk neutrality hypothesis, The relevant question is whether a theoretically sound economic model is able to provide a definition of risk capable of correctly pricing the forward premium.

In other words, although the two puzzles are similar with regards to the incapacity of traditional models to account for the risk premiums involved in these different markets, there is a non-shared characteristic: the predictability of returns based on interest rate differentials. It is possible that it was this specificity that lead researchers to adopt distinct agendas for investigating these puzzles. Engel (1996), for example, argues that, since this strong power of forward premium for forecasting exchange rate changes has no counterpart in the literature on equity returns, general equilibrium models are not likely to replicate this finding, there being no grounds to believe that the proposed solutions to puzzles in domestic financial markets can shed light on the FPP.

More recently, some works intending to better understand the behavior of foreign currency risk premiums have considered the possible relation between these puzzles, but none of them has provided empirical evidences toward this direction.

In this sense, our research agenda consists in showing this strong relation between these puzzles based on evidences that both empirical failures are related to the incapacity of the canonical CCAPM to provide a high volatile intertemporal marginal rate of substitution with reasonable values for the preferences parameters.

The first chapter (with João Victor Issler) has as a departure point that given free portfolio formation and the law of one price is equivalent to the existence of a SDF through Riesz representation theorem, which relies on less restrictive assumptions than those ones used in the CCAPM. Using two different purely statistical methodologies we extract time series for the SDF which are able to correctly price the excess market return
and also the excess returns that characterizes the FPP. Our results not only suggest that both puzzles are interwined, but also that American domestic are "representative" in the sense of characterizing a pricing kernel capable to price the foreign currency risk premium.

Because we do not spell out a full specified model it is hard to justify our calling the covariance of returns with the SDF as a risk measure. In the second chapter, we take the discussion of the previous chapter one step further by evaluating the performance of different models in pricing excess returns for each market. Once again, our goal is not to find a model that solves all the problems. Rather, our concern is to verify if the same failures and successes attained by the CCAPM in its various forms in pricing the excess returns of equity over short term risk-free bonds will be manifest in the case of forward exchange markets. Our main finding is that the same (however often unreasonable) values for the parameters are estimated for all models in both markets. In most cases, the rejections or otherwise of overidentifying restrictions occurs for the two markets, suggesting that success and failure stories for the equity premium repeat themselves in foreign exchange markets. Our results corroborate the findings in da Costa et al. (2006) that indicate a strong similarity between the behavior of excess returns in the two markets when modeled as risk premiums, providing empirical grounds to believe that the proposed preference-based solutions to puzzles in domestic financial markets can certainly shed light on the Forward Premium Puzzle.

However, even though one can write a successful risk premium model, which has been a surmountable task, there is a robust and uncomfortable empirical finding typical of the FPP that remains to be accommodated: domestic currency is expected to appreciate when domestic nominal interest rates exceed foreign interest rates.

In this sense, the third chapter (with Rodrigo de Losso) revisits these counterintuitive empirical findings working directly and only with the log-linearized Asset Pricing Equation. We are able to derive an equation that describes the currency depreciation movements based on a quite general framework and that has the conventional regression used in most empirical studies related to FPP as a particular case, therefore, being useful to identify the potential bias in the conventional regression due to a problem of omitted variable. We adopt a novel three-stage approach, wishing to analyse if this bias would be responsible for the disappointing findings reported in the literature. This chapter is still in process and the results are partially well succeed. In some tests we are not able to reject the null hypothesis that risk explains the FPP, while in other tests we reject the null.

## CHAPTER 1

# The Forward- and the Equity-Premium Puzzles: Two Symptoms of the Same Illness? ${ }^{1}$ 


#### Abstract

In this paper we revisit the question of the relationship between the equity and the forward premium puzzles. We construct atheoretical pricing kernels under the weak assumption that the law of one price holds and check whether the pricing kernel thus constructed prices correctly the equity and the forward premiums. We avoid the use of loglinearization by using the moments associated with the Euler equation to test the capacity of our pricing kernels to price correctly returns and excess returns. Our main finding, in our out of sample exercises, is that a pricing kernel constructed only from the information on American domestic assets is capable of accounting for both the equity and the forward premium puzzles. We fail to reject the null that the forward premium has zero price even when current and past values of the forward contract are used as instrument.


JEL Code: G12;G15
Keywords: Equity Premium Puzzle, Forward Premium Puzzle, Pricing Kernel, Common Feature, Principal Components Analysis.

## 1. Introduction

Puzzles are how economists denote systematic empirical violations of theoretically sound models. Puzzles play an important role in the advance of economic theory by creating research agendas aimed at either incorporating features capable of accommodating the empirical regularities that escape from the most stylized versions of these models, or, ultimately, overthrowing them. Two of the best known puzzles in economics are the equity premium - henceforth, EPP - and the forward premium puzzles - henceforth, FPP.

Despite some early attempts to treat the two puzzles in a unified framework - e.g., Mark (1985) — it was soon clear that, without an empirically successful model for pricing risk, as evidenced by the EPP, researchers would have no reason to believe that this would be the best approach to handle the FPP. In fact, the very question of whether the eventual development of a model capable of accommodating the EPP would allow for explaining the FPP appeared not to be well posed without actually having such model. This fact partly explains why the two puzzles gave birth to two separate research agendas: the first, within the scope of financial economics and the second mainly advanced by researchers in international economics. Also important was the existence of a specificity of the FPP with no parallel in the case of the EPP: the predictability of returns based on interest rate differentials. ${ }^{2}$

[^0]In this paper we revisit the two puzzles and ask whether they deserve distinct agendas, or whether they are but two symptoms of the same illness: the incapacity of our consumption based models to generate the implied behavior of $a$ pricing kernel ${ }^{3}$ that correctly prices all asset returns.

Our approach is indirect. If these two puzzles are solely a symptom of the inappropriateness of consumptionbased pricing kernels, then they will not be manifest when appropriate pricing kernels are used. If a single pricing kernel is capable of explaining not only the equity premium but also the forward premium then we have a good reason not to disperse our research effort on two completely different agendas, but rather to concentrate on a single one focused on rethinking consumption-based pricing kernels. If, on the contrary, we are not successful in finding a pricing kernel capable of explaining the equity and the forward premium, then, there may be a reason for these two literatures to evolve separately, since economic theory may have been rejected for different reasons, each specific to each problem.

A main issue of our approach is to find what would be appropriate pricing kernels. In an important paper, Hansen and Jagganathan (1991) have lead the profession towards return-based kernels instead of consumptionbased kernels, something the literature on factor and principal-component analysis in Finance concurs with e.g., Connor and Korajzcyk (1986), Chamberlain and Rothschild (1983), Bai (2005), and Araujo, Issler and Fernandes (2005). The whole idea is to combine statistical methods with economic theory to devise pricingkernel estimates that do not depend on preferences but solely on returns, thus being projections of stochastic discount factors - SDF - in the space of returns, at least to a first-order approximation.

We use two different model-free methodologies to extract a time series for the SDF as a function of the asset returns: i) a new methodology due to Araújo, Fernandes and Issler (2005) - AFI, henceforth - based on the fact that the (log of the) SDF is a common feature of asset prices, and ii) a statistical factor model. Applying both methodologies to two different data sets we estimate the SDF and show that the SDF compatible with the Equity premium behavior is also capable of accommodating the FPP.

Our first data set which we call global assets is comprised of assets traded in different countries. We are not able to reject the hypothesis that the SDF constructed from a factor model prices correctly all returns and excess returns on foreign assets as well as the excess return on equity over short term risk-free bonds in the US and the return on this risk-free bond. The same is not true for the other SDF, constructed with the AFI methodology, for which the Euler equation associated with the excess equity return is rejected for some of the instruments. Since foreign assets have been used to estimate the SDF, one may view our first exercise as being, in some sense, "in sample". Cochrane (2001) suggests that we should expect return-based kernels to appropriately price assets that are used in constructing them, but this information is not very useful, since this may be just a sign that there is in-sample over-fitting in constructing such estimates. The real test for any pricing kernel is not in sample, but out of sample, i.e., the ability to price assets not used in their construction or that are not highly correlated with those which are used in constructing them. So we take our second data set containing only US domestic assets.

When the SDF constructed from a traditional a factor model is used, we are not able to reject the null that excess returns have zero price for any currency: the SDF constructed from US domestic assets accounts for the FPP and, as one should expect, the EPP. In fact, in our out of sample exercises, the moments restrictions

[^1]associated with the Euler equations are not rejected for returns (and, consequently, for excess returns) for equities and operations with foreign assets for any of the instruments. In particular, the null is not rejected when the current and lagged values for the forward are used; one of the defining features of the FPP. When the SDF constructed with the AFI methodology is used, the result is not as stark. In general, we do not reject the null for excess returns but we do reject for returns. The SDF thus constructed seems to have too little volatility.

Overall, our results suggest that the two puzzles are indeed related and that the search for an economic model to account for the EPP is bound to double its prize by also accounting for the FPP.

The remainder of the paper is organized as follows. Section 2 gives an account of the literature that tries to explain the FPP, while in section 3 the techniques used to estimate the SDF are explained and the pricing tests are implemented. In section 4, the results are analyzed to understand the behavior of the SDF that solves FPP and EPP and then to draw parallels between them. The conclusion is in the last section.

## 2. Brief Literature Review

In this brief review we emphasize the research on the FPP that tries to explain it as a premium for accepting non-diversifiable risk.

The FPP, is how one calls the systematic violation of the "efficient-market hypothesis" for foreign exchange markets, where by "efficient-market hypothesis" we mean the proposition that the expected return to speculation in the forward foreign exchange market conditional on available information should be zero. ${ }^{4}$

Most studies report the FPP through the finding of $\hat{\alpha}_{1}<0$ when running the regression,

$$
\begin{equation*}
s_{\mathrm{t}+1}-s_{\mathrm{t}}=\alpha_{0}+\alpha_{1}\left(\mathrm{t} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+u_{\mathrm{t}+1} \tag{1}
\end{equation*}
$$

where $s_{\mathrm{t}}$ is the $\log$ of time $t$ exchange rate, $\mathrm{t} f_{\mathrm{t}+1}$ is the $\log$ of time $t$ forward exchange rate contract and $u_{\mathrm{t}+1}$ is the regression error. ${ }^{5}$ Notwithstanding the possible effect of Jensen inequality terms, testing the marketefficiency hypothesis of forward exchange rate market is equivalent to testing the null $\hat{\alpha}_{1}=1, \hat{\alpha}_{0}=0$, along with the uncorrelatedness of residuals from the estimated regression. The null is rejected in almost all studies.

As pointed out by Hansen and Hodrick (1980), however, this violation of market efficiency ought not to be viewed as evidence of market failure or some form of irrationality, since the uncovered parity need only hold exactly in a world of risk neutral agents or if the return on currency speculation is not risky. ${ }^{6}$ Nevertheless, the findings came to be called a puzzle due to, at least, one good reasons: the discrepancy from the null. Although one would not be surprised with $\hat{\alpha}_{1} \neq 1$, the fact that, $\hat{\alpha}_{1}<0$, in most studies - according to Froot (1990), the average value of $\hat{\alpha}_{1}$ is -0.88 for over 75 published estimates across various exchange rates and time periods - implies an expected domestic currency appreciation when domestic nominal interest rates exceed foreign interest rates. Moreover, the magnitude of the coefficients suggest a behavior of a risk premium which is hard to justify with our current models.

Despite these facts it is important to bear in mind that, without risk neutrality, rational expectations alone does not restrict the behavior of forward rates since, as suggested by Fama (1984): it should always be possible to include a risk premium with the right properties for reconciling the time series. The rejection of the null

[^2]only represents a true puzzle if risk may cannot explain the empirical regularities found in the data. The relevant question is, therefore, whether a theoretically sound economic model is able to provide a definition of risk capable of correctly pricing the forward premium. ${ }^{7}$

Once one recalls that a forward contract is a financial asset, the natural candidate for a theoretically sound model for pricing risk is the Consumption Capital Asset Pricing Model - CCAPM - of Lucas (1978) and Breeden (1979).

Assume that the economy has an infinitely lived representative consumer, whose preferences are representable by a von Neumann-Morgenstern utility function $u(\cdot)$. Then, the first order conditions for his optimal portfolio choice yields

$$
\begin{equation*}
1=\beta \mathbb{E}_{\mathrm{t}}\left[\frac{u^{\prime}\left(C_{\mathrm{t}+1}\right)}{u^{\prime}\left(C_{\mathrm{t}}\right)} R_{\mathrm{t}+1}^{\mathrm{i}}\right] \quad \forall i \tag{2}
\end{equation*}
$$

and, consequently,

$$
\begin{equation*}
0=\mathbb{E}_{\mathrm{t}}\left[\frac{u^{\prime}\left(C_{\mathrm{t}+1}\right)}{u^{\prime}\left(C_{\mathrm{t}}\right)}\left(R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}\right)\right] \quad \forall i, j \tag{3}
\end{equation*}
$$

where $\beta \in(0,1)$ is the discount factor in the representative agent's utility function, $\mathbb{E}_{\mathrm{t}}(\cdot)$ denotes the conditional expectation given the information available at time $t, R_{\mathrm{t}+1}^{\mathrm{i}}$ and $R_{\mathrm{t}+1}^{\mathrm{j}}$ denote, respectively, the real gross return on assets $i$ and $j$ at time $t+1$ and, $C_{\mathrm{t}}$ is aggregate consumption at time $t$.

Under the CCAPM, Euler equations (2) and (3) should hold for all assets and portfolios and at all times. Moreover, it is apparent from (3) that the expected excess returns are zero when deflated by the representative agent's intertemporal marginal rate of substitution, $u^{\prime}\left(C_{\mathrm{t}+1}\right) / u^{\prime}\left(C_{\mathrm{t}}\right)$ : the relevant measure of risk.

Define the covered, $R^{\mathrm{C}}$, and the uncovered return, $R^{\mathrm{U}}$, on trading of foreign government bonds as

$$
\begin{equation*}
R_{\mathrm{t}+1}^{\mathrm{C}}=\frac{\mathrm{t} F_{\mathrm{t}+1}\left(1+i_{\mathrm{t}}^{*}\right) P_{\mathrm{t}}}{S_{\mathrm{t}} P_{\mathrm{t}+1}} \quad \text { and } \quad R_{\mathrm{t}+1}^{\mathrm{U}}=\frac{S_{\mathrm{t}+1}\left(1+i_{\mathrm{t}}^{*}\right) P_{\mathrm{t}}}{S_{\mathrm{t}} P_{\mathrm{t}+1}} \tag{4}
\end{equation*}
$$

where ${ }_{\mathrm{t}} F_{\mathrm{t}+1}$ and $S_{\mathrm{t}}$ are the forward and spot prices of foreign currency in terms of domestic currency, $P_{\mathrm{t}}$ is the dollar price level and $i_{\mathrm{t}}^{*}$ represents nominal net return on a foreign asset in terms of the foreign investor's preferences.

Next substitute $R^{\mathrm{C}}$ for $R^{\mathrm{i}}$ and $R^{\mathrm{U}}$ for $R^{\mathrm{j}}$ in (3), to get

$$
\begin{equation*}
0=\mathbb{E}_{\mathrm{t}}\left\{\frac{u^{\prime}\left(C_{\mathrm{t}+1}\right)}{u^{\prime}\left(C_{\mathrm{t}}\right)} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)\left[\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}\right]}{S_{\mathrm{t}} P_{\mathrm{t}+1}}\right\} \tag{5}
\end{equation*}
$$

Equation (5) states that the risk adjusted forward premium, should be zero.
As far as we know, Mark (1985) was the first to test (5). Assuming that preferences exhibit constant relative risk aversion, he estimates the parameter $\gamma$ in $C^{1-} /(1-\gamma)$ using Hansen's (1982) Generalized Method of Moments (GMM). He then tests the overidentifying restrictions to access the validity of the model. He reports a coefficient of relative risk aversion, $\gamma$, above 40 , well above what one considers a reasonable value, ${ }^{8}$ and rejects the overidentifying restrictions when the forward premium and its lags are used as instruments. Similar results are reported later by Modjtahedi (1991); Later, with a different, larger data set, Hodrick (1989) report values of the coefficient of relative risk aversion, $\gamma$, above 60 , but does not reject the over-identifying restrictions, while

[^3]Engel (1996) reports some estimates for $\gamma$ in excess of 100. A more recent attempt to use Euler equations to account for the FPP is Lustig and Verdelhan (2006) which use a more general utility function, but once again finds values of risk aversion above 100. It is apparent from these studies that we cannot account for the FPP with reasonable parameters of risk aversion.

Are we, however, to be surprised with these findings? If we recall that the EPP is identified with the failure to explain the excess return of equity over risk free short term bonds $-R_{\mathrm{t}+1}^{\mathrm{i}}=\left(1+i_{\mathrm{t}}^{\mathrm{SP}}\right) P_{\mathrm{t}} / P_{\mathrm{t}+1}$ and $R^{\mathrm{j}}=\left(1+i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}} / P_{\mathrm{t}+1}$, in (3), where $i_{\mathrm{t}+1}^{\mathrm{SP}}$ is nominal return on $\mathrm{S} \& \mathrm{P} 500$ and $i_{\mathrm{t}+1}^{\mathrm{b}}$ nominal return on the U.S. Treasury Bill - with reasonable parameters of risk aversion for $(3),{ }^{9}$ why should we expect the same model to account for the FPP? This path for explaining the behavior of the forward premium seems to be doomed from its start!

The lack of a reliable model to account for risk was in great deal responsible for the separation of the research agendas involving the two puzzles. Research regarding the FPP, is mostly done within the scope of international economics, like in Fama and Farber (1979), Hodrick (1981) and Lucas (1982). It emphasizes international affine term structure models and/or a microstructure approach. Research involving the EPP, on the other hand, has concentrated on adding to (3) additional state variables, as Constantinides and Duffie (1996) and Campbell and Cochrane (1999).

There is, however, a second characteristic of the FPP that may have played a role: the predictability of returns on currency speculation. Because $\hat{\alpha}_{1}<0$ and given that the auto-correlation of risk premium is very persistent, interest rates differential predict excess returns. Although predictability in equity markets has also been extensively documented, this is not a defining feature of the EPP. It is, however, a defining feature of
to guarantee, through Riesz representation theorem, the existence of a SDF. The validity of the law of one price is, of course, a much weaker assumption than those that underlie the CCAPM.

Our strategy consists in estimating an SDF using purely statistical methodologies and evaluating whether this SDF accounts for both puzzles.

There is no claim of originality in our choice of not using consumption data. Hansen and Hodrick (1983), Hodrick and Srivastava (1984), Cumby (1988), Huang (1989) and Lewis (1990) all implemented latent variable models that avoid the need of specifying a model for the pricing kernel by treating the return on a benchmark portfolio as a latent variable. They use the fact that movements in assets expected return should be proportional to movements in the expected return on a benchmark portfolio. Their results met with partial success: all these papers reject the unbiasedness hypothesis but are in conflict with each other with regards to the rejection of restrictions imposed by the latent variable model. This line of research does not try to relate the EPP and the FPP, not making use of the large cross-section of returns in US domestic markets.

More closely related to our work is Korajczyk and Viallet (1992). Applying the arbitrage pricing theory APT - to a large set of assets from many countries they test whether including the factors as the prices of risk reduces the predictive power of the forward premium. They are able to show that the forward premium has lower though not zero predictive power when one includes the factors. Their results should be take with some caution since they also show that the inclusion of factors often reduces the overall explaining power (as measured by the $R^{2}$ ) of the model. Their model work differ from ours in several dimensions, most notably, they do not try to relate the two puzzles, and they do not perform any out of sample exercises.

Also worth mentioning is Backus et al. (1995) which poses the question of whether a pricing kernel can be found that satisfies, at the same time, (log linearized versions of) equations

$$
\begin{equation*}
0=\mathbb{E}_{\mathrm{t}}\left\{M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}}\right\} \tag{8}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{1}{\mathbb{E}_{\mathrm{t}}\left(M_{\mathrm{t}+1}\right)}=R_{\mathrm{t}+1}^{\mathrm{f}} \tag{9}
\end{equation*}
$$

where $R_{t+1}^{f}$ is the risk-free rate of return. Equation (8) is derived from (7) and (4).
In our in sample exercise, we explore a large panel of assets returns to extract a SDF with purely statistical methods. Next we check whether the SDF thus constructed prices correctly both the return and the excess return on equities and on trading of foreign government bonds, which means that the nature of the question implicitly answered is similar to the one posed by Backus et al. (1995), albeit adding a pricing test of the excess return on equity over risk free short term bonds in the US, ${ }^{11}$ i.e., if the equation

$$
\begin{equation*}
0=\mathbb{E}_{\mathrm{t}}\left\{M_{\mathrm{t}+1} \frac{\left(i_{\mathrm{t}+1}^{\mathrm{SP}}-i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}}\right\} \tag{10}
\end{equation*}
$$

is satisfied.
In our out of sample exercise, however, we go a step further. We use only data from the US domestic market and construct an SDF which prices correctly the equity premium. Next, we take this SDF, which, we must

[^4]emphasize, was not built to satisfy (8), and verify if it accounts for the FPP. To anticipate our results, we cannot reject (8), for any of the instruments, that including the current and lagged forward premium.

Although the solution for the FPP and the EPP can be characterized by the search of a model able to satisfy (8) and (10) respectively, the asset pricing literature supposes that this successful model should also be able to satisfy (6) for the returns directly. With this in mind and intending to raise the power of our tests, we also consider pricing returns on the covered trading of foreign government bonds and on the risk-free short term US bond, testing respectively the equations

$$
\begin{align*}
& 1=\mathbb{E}_{\mathrm{t}}\left\{M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}}{S_{\mathrm{t}}}\right\} \text { and }  \tag{11}\\
& 1=\mathbb{E}_{\mathrm{t}}\left\{M_{\mathrm{t}+1} \frac{\left(1+i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}}\right\} . \tag{12}
\end{align*}
$$

Another novelty of our procedure is to combine an atheoretical construct within the Euler equation framework. Almost all research that attempts to explain the FPP by including a risk premium adopts log-linearizations that allow equation (8) to be written as

$$
\begin{equation*}
s_{\mathrm{t}+1}-s_{\mathrm{t}}=-\left[\frac{1}{2} \mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)+\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)\right]+\left(\mathrm{t} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+\varepsilon_{\mathrm{t}+1} \tag{13}
\end{equation*}
$$

where $\mathbb{V}_{\mathrm{t}}(\cdot)$ and $\operatorname{cov}_{\mathrm{t}}(\cdot, \cdot)$ denote, respectively, the conditional variance and covariance given the available information at time $t, \pi_{\mathrm{t}+1}$ is the (log of) domestic price variation, $m_{\mathrm{t}+1}$ is the $\log$ of the pricing kernel and $\varepsilon_{\mathrm{t}+1}$ denotes the innovation in predicting $\ln \left(M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{U}}\right)-\ln \left(M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{C}}\right)$.

The use of such linearization makes the comparison with equation (1) immediate, which may explain the preference of researchers in the field. Equation (13) describes the movement of currency depreciation and is useful in drawing implications for the SDF and in deriving empirical tests of FPP. ${ }^{12}$

Comparing (1) with (13), it is apparent that the latter is a particular case of the former, where term $\frac{1}{2} \mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)+\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)$ is assumed to be time-invariant. The $\mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)$ term , which we referred to before as the Jensen's inequality term, appears only because we defined the expected rate of depreciation in logarithms. According to Bekaert and Hodrick (1993) omitting this term from regression (1) is not responsible for finding $\hat{\alpha}_{1}<0$. Therefore, all the action must come from the term $\operatorname{cov}_{\mathbf{t}}\left(s_{\mathbf{t}+1}, m_{\mathbf{t}+1}-\pi_{\mathbf{t}+1}\right)$. This term must exhibit considerable variation if one is to accommodate the evidence regarding the forward premium. Given that asset returns have clear signs of conditional heteroskedasticity, ${ }^{13}$ time invariance for the covariance term does not seem to be a sound assumption, leading one to wonder whether this might not be the key to accounting for $\hat{\alpha}_{1}<0$.

Log-linearization is clearly an unnecessary detour. As we have seen, the puzzle is manifest in logs and in levels. Moreover, to go from (8) to (13) one must add some bold assumptions on the joint distribution of assets and the SDF that may not be verified in many data sets. In particular one must impose the condition that $M_{\mathrm{t}+1}$ is positive. Although the existence of a strictly positive SDF can be proved under no arbitrage, what

[^5]we must impose is $M_{\mathrm{t}+1}^{*}>0$, where $M_{\mathrm{t}+1}^{*}$ is the unique SDF in the span of traded assets, labeled mimicking portfolio. When markets are complete, $M_{\mathrm{t}+1}^{*}=M_{\mathrm{t}+1}$ and no-arbitrage is all that we need. With incomplete markets, however there is a continuum of SDF's pricing all traded securities. Each $M_{\mathrm{t}+1}$ may be written as $M_{\mathrm{t}+1}=M_{\mathrm{t}+1}^{*}+\nu_{\mathrm{t}+1}$ for some $\nu_{\mathrm{t}+1}$ with $\mathbb{E}_{\mathrm{t}}\left[\nu_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}\right]=0 \forall i$. Although there is an $M_{\mathrm{t}+1}>0$, one may not guarantee that $M_{\mathrm{t}+1}^{*}>0$. By working directly with (7) we avoid dealing with these issues, while keeping the possibility of testing the conditional moments through the use of lagged instruments, along the lines of Hansen and Singleton (1983) and Mark (1985).

## 3. Econometric Tests

In this section, we show how to implement direct pricing tests using the forward- and the equity-premium, in an Euler equation framework. To be clear about this issue, we keep the notation $M_{\mathrm{t}+1}$ for the pricing kernel, without any specific functional forms depending on preferences and consequently on consumption.

### 3.1. Pricing Test

Applying the Law-of-Iterated Expectations to (8), (11), (10) and (12) we get, respectively:

$$
\begin{align*}
& 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}}\right\}  \tag{14}\\
& 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t}}{} \frac{F_{\mathrm{t}+1}}{S_{\mathrm{t}}}-1\right\}  \tag{15}\\
& 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{\left(i_{\mathrm{t}+1}^{\mathrm{SP}}-i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}}\right\} \text { and }  \tag{16}\\
& 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{\left(1+i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}}-1\right\} . \tag{17}
\end{align*}
$$

Equations (14) and (15) produce testable implications of the forward premium using an estimator of $M_{\mathrm{t}+1}$, where

$$
M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t}}{} \frac{F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}} \quad \text { and } \quad M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}}{S_{\mathrm{t}}}-1
$$

must both have zero means. In order to test the equity premium, the equations (16) and (17) are useful, where

$$
M_{\mathrm{t}+1} \frac{\left(i_{\mathrm{t}+1}^{\mathrm{sP}}-i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}} \quad \text { and } \quad M_{\mathrm{t}+1} \frac{\left(1+i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}}-1
$$

must both have zero mean.
Zero-mean tests such as those in (14), (15), (16) and (17) are straightforward to implement. The only issue is how to construct a robust estimate for the variance of sample-mean estimates, taking into account possible serial correlation and heteroskedasticity in their components. Here we employ the non-parametric estimate proposed by Newey and West (1987).

These are not, however, the only testable implications of (8), (11), (10) and (12). Consider $z_{\mathrm{t}}$ to be any instrumental variable observed up to time $t$, therefore measurable with respect to $\mathbb{E}_{t}(\cdot)$. We may also employ scaled returns and scaled excess-returns - defined as $R_{\mathrm{t}+1}^{\mathrm{i}} \times z_{\mathrm{t}}$ and $\left(R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}\right) \times z_{\mathrm{t}}$, respectively - to test the conditional moments and consequently to derive some extra implications from the presence of information,
which is particularly important for the FPP since, when the CCAPM is employed, the over-identifying restriction associated with having the forward premium as an instrument is usually rejected: a manifestation of its predictive power.

Multiply now (8) and (11) by $z_{\mathrm{t}}$ and apply once again the law-of-iterated expectations to get:

$$
\begin{align*}
& 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}} \times z_{\mathrm{t}}\right\}, \text { and }  \tag{18}\\
& 0=\mathbb{E}\left\{\left[M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}}{S_{\mathrm{t}}}-1\right] \times z_{\mathrm{t}}\right\} \tag{19}
\end{align*}
$$

Following this same procedure to the equity premium, we get:

$$
\begin{align*}
& 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{\left(i_{\mathrm{t}+1}^{\mathrm{S}}-i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}} \times z_{\mathrm{t}}\right\} \text { and }  \tag{20}\\
& 0=\mathbb{E}\left\{\left[M_{\mathrm{t}+1} \frac{\left(1+i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}}-1\right] \times z_{\mathrm{t}}\right\} \tag{21}
\end{align*}
$$

The equations (18) and (19) are then defined as the null hypothesis testing the FPP, while (20) and (21) as the null hypotheses testing the EPP, where we can consider the unconditional test provided by the equations $(14),(15),(16)$ and (17) as a special case for $z_{\mathrm{t}}=1$.

### 3.2. Return-Based Pricing Kernels

In this section we provide a brief description of return-based estimates of the SDF or pricing kernel. There are two basic techniques employed here. The first views the mimicking portfolio (or pricing kernel) as a common feature of asset returns up to a first-order logarithmic approximation. It has its roots in the work of Hansen and Singleton (1983), exploited partially in Engle and Marcucci (2006) and to the fullest in Araujo, Issler and Fernandes (2005). The second uses principal-component and factor analyses to extract common components of asset returns. It can be traced back to the work of Ross (1976), developed further by Chamberlain and Rothschild (1983), and Connor and Korajczyk (1986, 1993). Additional references are Hansen and Jagganathan (1991) and Bai (2005).

The basic idea behind return-based pricing kernels is that asset prices (or returns) convey information about the intertemporal marginal rate of substitution in consumption. In equilibrium, all returns must have a common feature (factor), which can be removed by subtracting any two returns. This common feature is the SDF. One way to see this, is to realize that although

$$
\begin{aligned}
\mathbb{E}\left\{M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}\right\} & =1, \text { for all } i=1,2, \cdots, N, \\
\mathbb{E}\left\{M_{\mathrm{t}+1}\left(R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}\right)\right\} & =0, \text { for all } i \neq j .
\end{aligned}
$$

Hence, even though $R_{\mathrm{t}+1}^{\mathrm{i}}$ is not orthogonal to $M_{\mathrm{t}+1}$ itself, if we subtract the returns of any two assets, their return differential, $R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}$, will be orthogonal to $M_{\mathrm{t}+1}$. Hence, $R_{\mathrm{t}+1}^{\mathrm{i}}$ contains the feature $M_{\mathrm{t}+1}$ but $R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}$ does not.

Because every asset return $R_{\mathrm{t}+1}^{\mathrm{i}}$ contains "a piece" of $M_{\mathrm{t}+1}$, if we combine a large number of returns, the average idiosyncratic component of returns will vanish in limit. Hence, if we choose our weights properly, we
may end up with the common component of returns, i.e., the SDF. Of course, all these methods are asymptotic, either $N \rightarrow \infty$ or $N, T \rightarrow \infty$, and rely on weak-laws of large numbers to provide consistent estimators of the mimicking portfolio - the systematic portion of asset returns.

Multifactor Models Factor models summarize the systematic variation of the $N$ elements of the vector $\mathbf{R}_{\mathrm{t}}=\left(R_{\mathrm{t}}^{1}, R_{\mathrm{t}}^{2}, \ldots, R_{\mathrm{t}}^{\mathrm{N}}\right)^{\prime}$ using a reduced number of $K$ factors, $K<N$. Consider a $K$-factor model in $R_{\mathrm{t}}^{\mathrm{i}}$ :

$$
\begin{equation*}
R_{\mathrm{t}}^{\mathrm{i}}=a_{\mathrm{i}}+\sum_{\mathrm{k}=1}^{\mathrm{K}} \beta_{\mathrm{i} ; \mathrm{k}} f_{\mathrm{k} ; \mathrm{t}}+\eta_{\mathrm{it}} \tag{22}
\end{equation*}
$$

where $f_{\mathrm{k} ; \mathrm{t}}$ are zero-mean pervasive factors and, as is usual in factor analysis,

$$
\operatorname{plim} \frac{1}{N} \sum_{\mathrm{i}=1}^{\mathrm{N}} \eta_{\mathrm{i} ; \mathrm{t}}=0
$$

Denote by $\Sigma_{\mathbf{r}}=\mathbb{E}\left(\mathbf{R}_{\mathrm{t}} \mathbf{R}_{\mathrm{t}}^{\prime}\right)-\mathbb{E}\left(\mathbf{R}_{\mathrm{t}}\right) \mathbb{E}\left(\mathbf{R}_{\mathrm{t}}^{\prime}\right)$ the variance-covariance matrix of returns. The first principal component of the elements of $\mathbf{R}_{\mathrm{t}}$ is a linear combination $\theta^{\prime} \mathbf{R}_{\mathrm{t}}$

Given these coefficients, can easily get an estimate of $M_{\mathrm{t}}^{*}$. The relation between $(\lambda, \gamma)$ and $(a, b)$, given by

$$
a \equiv \frac{1}{\gamma} \text { and } b \equiv-\gamma\left[\operatorname{cov}\left(f f^{\prime}\right)\right]^{-1} \lambda
$$

allows us to ensure the equivalence between this beta pricing model and the linear model for the SDF, i.e. that

$$
\widetilde{M_{\mathrm{t}}^{*}}=a+\sum_{\mathrm{k}=1}^{\mathrm{K}} b_{\mathrm{k}} f_{\mathrm{k} ; \mathrm{t}} \text { and } \mathbb{E}\left(\widetilde{M_{\mathrm{t}}^{*}} R^{\mathrm{i}}\right)=1, i=1,2, \ldots, N
$$

hold.
The number of factors used in the empirical analysis is an important issue. We expect $K$ to be rather small, but have some flexibility for this choice. We took a pragmatic view here, increasing $K$ until the estimate of $M_{\mathrm{t}}^{*}$ changed very little due to the last increment; see for instance Lehmann and Modest (1988) and Connor and Korajczyk (1988). We also performed a robustness analysis for the results of all of our statistical tests using different estimates of $M_{\mathrm{t}}^{*}$ associated with different $K$ 's. Results changed very little around our choice of $K$. Our final choice of $K$ is in accordance with Connor and Korajczyk (1993), who examined returns from stocks listed on the New York Stock Exchange and the American Stock Exchange.

AFI Methodology In a recent paper, Araujo, Issler and Fernandes proposed an alternative return-based estimates of the SDF using panel-data asymptotics. After using an exact Taylor expansion of the Pricing Equation, they are able to express the logarithm of returns $\left(r_{\mathrm{t}}^{\mathrm{i}}\right)$ as a function of the logarithm of the mimicking portfolio $\left(m_{\mathrm{t}}^{*}\right)$ :

$$
\begin{equation*}
r_{\mathrm{t}}^{\mathrm{i}}=-m_{\mathrm{t}}^{*}-\gamma_{\mathrm{i}}^{2}+\varepsilon_{\mathrm{i} ; \mathrm{t}}-\sum_{\mathrm{j}=1}^{\infty} b_{\mathrm{i} ; \mathrm{j}} \varepsilon_{\mathrm{i} ; \mathrm{t}-\mathrm{j}}, \quad i=1,2, \ldots, N \tag{23}
\end{equation*}
$$

where

$$
\begin{equation*}
\varepsilon_{\mathrm{i} ; \mathrm{t}}=\left[m_{\mathrm{t}}^{*}-\mathbb{E}_{\mathrm{t}-1}\left(m_{\mathrm{t}}^{*}\right)\right]+\left[r_{\mathrm{t}}^{\mathrm{i}}-\mathbb{E}_{\mathrm{t}-1}\left(r_{\mathrm{t}}^{\mathrm{i}}\right)\right]=q_{\mathrm{t}}+v_{\mathrm{i} ; \mathrm{t}} \tag{24}
\end{equation*}
$$

and the last equality defines notation. Under a set of plausible assumptions, they consider a cross-sectional average of (23), showing that $M_{\mathrm{t}}^{*}$, can be consistently estimated as $N, T \rightarrow \infty$, using

$$
\begin{equation*}
\widehat{M_{\mathrm{t}}^{*}}=\frac{\bar{R}_{\mathrm{t}}^{\mathrm{G}}}{\frac{1}{\mathrm{~T}} \sum_{\mathrm{j}=1}^{\top}\left(\bar{R}_{\mathrm{j}}^{\mathrm{G}} \bar{R}_{\mathrm{j}}^{\mathrm{A}}\right)} \tag{25}
\end{equation*}
$$

where $\bar{R}_{\mathrm{t}}^{\mathrm{G}}=\prod_{\mathrm{i}=1}^{\mathrm{N}}\left(R_{\mathrm{t}}^{\mathrm{i}}\right)^{-1=\mathrm{N}}$ and $\bar{R}_{\mathrm{t}}^{\mathrm{A}}=\frac{1}{\mathrm{~N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} R_{\mathrm{t}}^{\mathrm{i}}$ are respectively the geometric average of the reciprocal of all asset returns and the arithmetic average of all asset returns. It is important to bear in mind that there is no need to assume complete markets. However, if markets are complete, the realization of the SDF at time $t$, denoted by $M_{\mathrm{t}}$, can be consistently estimated using (25), while, if markets are incomplete, the mimicking portfolio will be consistently estimated.

## 4. Empirical Results

### 4.1. Data and Summary Statistics

In principle, whenever econometric or statistical tests are performed, it is preferable to employ a long data set either in the time-series or in the cross-sectional dimension. There are some obvious limitations, especially
when the FPP is concerned, the main one being due to the forward rate series, since a sufficiently large timeseries is hardly available ${ }^{15}$. Regarding the EPP these limitations are less severe. In order to have a common sample we covered the period starting in 1977:1 and ending in 2004:3, with quarterly frequency. We collected foreign exchange data for the following countries: Canada, Germany, Japan, Switzerland, U.K. and the U.S.

When testing the EPP or the FPP we need first to compute excess returns $R_{\mathrm{t}}^{\mathrm{i}}-R_{\mathrm{t}}^{\mathrm{j}}$. On that regard, our data set is composed of the following. To study the EPP we used the US $\$$ real returns on the S\&P500 and on 90-day T-Bill. To analyze the FPP, we used the US\$ real returns on short-term British, Canadian, German, Swiss and Japanese government bonds, where both spot and forward exchange-rate data were used to transform returns denominated in foreign currency into US\$. The forward rate series were extracted from the Chicago Mercantile of Exchange database, while the spot rate series were extracted from Bank of England database.

A second ingredient for testing these two puzzles is the construction of return-based estimates of the pricing kernel $\left(M_{\mathrm{t}}^{*}\right)$. For available data sets, there is a trade-off between $N$ and $T$-in order to get a larger $N$, one must accept a reduction in $T$. For that reason, we used two alternative data sets. The first is what we have labelled a global data set, mostly composed of aggregate returns on stocks and government bonds for G7 countries. It covers US\$ real returns on G7-country stock indices and short-term government bonds, where spot exchange rate data were used to transform returns denominated in foreign currency into US $\$$ and the consumer price index of services and nondurable goods in the US was used as a deflator. In addition to G7 returns on stocks and bonds, it also contains US\$ real returns on gold, US real estate, bonds on AAA US corporations, Nasdaq, Dow Jones and the S\&P500. The US government bond is chosen to be the 90-day T-Bill. Regarding data sources, the returns on G7 government bonds were extracted from IFS/IMF, the returns on Nasdaq and Dow Jones Composite Index were extracted from Yahoo finance, the returns on real-estate trusts were extracted from the National Association of Real-Estate Investment Trusts in the US ${ }^{16}$, while the remaining series were extracted from the DRI database. Our sample period starts in 1977:1 and ends in 2001:3, with quarterly frequency. These portfolios of assets cover a wide spectrum of investment opportunities across the globe, which is an important element of our choice of assets, since diversification is recommended for both methods of computing $M_{\mathrm{t}}^{*}$.

The second data set used to compute $M_{\mathrm{t}}^{*}$ includes only assets that are traded in the U.S. In this case, diversification is not exercised to the fullest, since several investment opportunities abroad were not included in computing the SDF. On the other hand, the data set is much more disaggregated than the previous one, being comprised of US\$ real returns on a short-term U.S. government bond and on two hundred U.S. stocks having the largest transaction volume from 1990:1 to 2004:3, on a quarterly frequency. Data were extracted from the CRSP database.

In terms of the notation used in the tables below, we adopt the following: $M_{\mathrm{t}}^{*}$ estimates using the commonfeature technique of Araujo, Issler and Fernandes are labelled $\widehat{M_{\mathrm{t}}}$, while $M_{\mathrm{t}}^{*}$ estimates using factor models are labelled $\widetilde{M_{\mathrm{t}}}$. Estimates constructed using the "global" database of assets use a superscript "GL," while estimates constructed using the "U.S." database use a superscript "US." Therefore, $\widehat{M_{\mathrm{t}} \mathrm{GL}}$ denotes the SDF estimator using the technique in Araujo, Issler and Fernandes and the global database of assets, while $\widetilde{M_{\mathrm{t}}}$ denotes the SDF estimator using factor models and the U.S. database of assets, and so on.

[^6]We employ the following additional variables in orthogonality tests: real consumption instantaneous growth rates, real GDP instantaneous growth rates, and the logarithm of the ratio between real consumption and real GDP. We also employ additional financial variables that are specific to each test performed, and are listed in the appropriate tables of results. All macroeconomic variables were extracted from FED's FRED database.

Table 1 presents summary statistics of our database over the period 1977:1 to 2004:3. The average real return on the $\mathrm{S} \& \mathrm{P} 500$ is $8.67 \%$ at an annual rate, while that of the 90 -day T-Bill is $1.76 \%$. Real stock returns are much more volatile than the US Treasury Bill return - annualized standard deviation of $16.02 \%$ and $1.60 \%$ respectively. Over the same period, except for the Swiss case, the real return on covered trading of foreign bonds show means and standard deviations quite similar to that of the U.S. Treasury Bill. Regarding the return on uncovered trading, means range from $2.47 \%$ to $4.70 \%$, while standard deviations range from $5.44 \%$ to $16.18 \%$.

We observe a Sharp ratio for the U.S. stock market of 0.44 , while the Sharp ratio of the uncovered trading of foreign bonds ranges from 0.04 to 0.28 . According to Shiller (1982), Hansen and Jagganathan (1991) and Cochrane and Hansen (1992), an extremely volatile SDF is required to match the high equity Sharpe ratio of the U.S. Hence, the smoothness of aggregate consumption growth is the main reasons behind the EPP, but could not be so for the FPP. However, since the Sharp ratios of foreign bonds are lower, this may be an evidence that a SDF that prices correctly the equity premium, will also price correctly the forward premium, while the converse may not be true.

### 4.2. SDF Estimates

We employ two data sets to estimate $M_{\mathrm{t}}$. The first makes heavily use of a reduced number of global portfolios. Despite this reduced number, it is plausible to expect the operation of a weak-law of large numbers, since these portfolios contain many assets. For example, we use aggregate returns on equity for G7 countries. For each country, these portfolios contain hundreds of asset (sometimes more than a thousand, depending on the country being considered). Therefore, the final average containing these portfolios will be formed using thousands of assets worldwide. Of course, it would be preferable to work at a more disaggregated level, e.g., at the firm level. However, there is no comprehensive database with a long enough time span containing worldwide returns at the firm level. As mentioned before, there is a trade-off between $N$ and $T$ for available databases, which today limits empirical studies.

The second database used here to estimate $M_{\mathrm{t}}$ is comprised of two hundred U.S. stocks (those with the 200 largest volumes) and also of a short-term U.S. government bond. Therefore, it is completely U.S. based available at a very disaggregated level - mostly firm returns. Unfortunately, the cost of gathering similar data sets for other G7 countries is too big, explaining why we did not attempt to do it here.

Estimates $\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}, \widetilde{M_{\mathrm{t}}^{\mathrm{GL}}}, \widehat{M_{\mathrm{t}}^{\mathrm{US}}}$ and $\widetilde{M_{\mathrm{t}}^{\mathrm{US}}}$ are plotted in Figure 1, which also includes their summary statistics. As we should expect, their means are slightly below unity in all cases. Moreover, factor-model estimates are much more volatile than estimates using the common features technique, with standard deviations at least 3 times larger. Nevertheless, SDF estimates are highly correlated: the correlation coefficient between $\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}$ and $\widetilde{M_{\mathrm{t}}^{G L}}$ is 0.740 while that between $\widehat{M_{\mathrm{t}}^{\mathrm{US}}}$ and $\widetilde{M_{\mathrm{t}}^{\mathrm{US}}}$ is 0.799 . SDF estimates based solely on U.S. assets are also more volatile than those based on a "global market," with volatility about twice as large.

Since we want the pricing exercise to be out-of-sample, i.e., estimates of $M_{\mathrm{t}}$ must price assets not used in computing it, or not highly correlated with those which are, we carefully changed the asset list used in
estimating $M_{\mathrm{t}}$ in each test performed. In doing so, we avoid the in-sample over-fitting criticism in Cochrane (2001) applied to SDF estimates. Of course, when we are pricing assets on G7 countries other than the U.S., using SDF estimates based solely on U.S. assets, we are performing out-of-sample pricing. However, when we use global estimates we take the following precaution: if, for example, we are pricing the excess return of the uncovered over the covered trading of the British government bond, we exclude from the list of assets used to compute the SDF the short-term British government bond and the British stock index. We do the same for all the other countries as well, therefore using a different SDF estimate for each country.

Finally, when the multi-factor model is used, since factors may be viewed as traded portfolios, it is interesting to examine which assets carry the largest weights in the SDF composition. For the global data set, in order to specify the three factors used, the assets with largest weights are the German, British and American stock indices, the German, Japanese and British government bonds and Nasdaq. For the U.S. based data set, where there are up to six pervasive factors, the most relevant stocks are: Informix corporation (13th largest volume), AMR corporation DEL (64th), Emulex corporation (98th), Ericsson L M Telephone corporation (99th), Iomega corporation (118th), LSI corporation (124th), Lam Resch corporation (125th), Advanced Micro Services inc. (154th) and 3-Com corporation (193th).

### 4.3. Pricing-Test Results

The results of the FPP tests provided by (18) and (19):

$$
\begin{align*}
& H_{0}: \quad 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}} \times z_{\mathrm{t}}\right\} \text {, and }  \tag{26}\\
& H_{0} \quad: \quad 0=\mathbb{E}\left\{\left[M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}}{S_{\mathrm{t}}}-1\right] \times z_{\mathrm{t}}\right\} \tag{27}
\end{align*}
$$

are reported in Tables 2 to 6 when $\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}$ and $\widetilde{M_{\mathrm{t}}^{\mathrm{GL}}}$ are used, and are reported in Tables 8 to 12 when $\widehat{M_{\mathrm{t}}^{\mathrm{US}}}$ and $\widetilde{M_{\mathrm{t}}^{U S}}$ are used.

Tables 7 and 13 present the results of the following EPP tests (20) and (21):

$$
\begin{align*}
& 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{\left(i_{\mathrm{t}+1}^{\mathrm{SP}}-i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}} \times z_{\mathrm{t}}\right\} \text { and }  \tag{28}\\
& 0=\mathbb{E}\left\{\left[M_{\mathrm{t}+1} \frac{\left(1+i_{\mathrm{t}}^{\mathrm{b}}\right) P_{\mathrm{t}}}{P_{\mathrm{t}+1}}-1\right] \times z_{\mathrm{t}}\right\} \tag{29}
\end{align*}
$$

when $\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}$ and $\widetilde{M_{\mathrm{t}}^{\mathrm{GL}}}$ are used, and when $\widehat{M_{\mathrm{t}}^{\mathrm{US}}}$ and $\widetilde{M_{\mathrm{t}}^{U S}}$ are used, respectively.
For Tables 2 to 6 , and 8 to 12 , the first column reports the result of testing the null hypothesis in the unconditional test, while the remaining columns report the results of testing conditionally the null hypothesis in (26) for a variety of scales (instruments). Hypotheses testing used robust standard-error estimates whenever the data showed signs of either heteroskedasticity or serial correlation, which were tested at the $5 \%$ level using the ARCH LM test and the Breusch-Godfrey LM test respectively. For Tables 7 and 13 the same applies regarding (28).

For Tables 2 to 7 , when $\widetilde{M_{\mathrm{t}}^{\mathrm{GL}}}$ is used, there are no signs of either the FPP or the EPP for all countries. When $\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}$ is used, there some signs of the EPP but not of the FPP. Since $\mathbb{E}\left\{\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}\right\}$ is very close to $\mathbb{E}\left\{\widetilde{M_{\mathrm{t}}^{\mathrm{GL}}}\right\}$,
this orthogonality failure is probably related to the covariance between $\widehat{M_{\mathrm{t}} \mathrm{GL}}$ and the scaled excess return of S\&P500 over the US short term bond ${ }^{17}$. Although all these tests used a set of global assets, SDF estimates varied from country to country, in order to exclude the possibility of in-sample over-fitting and the criticism in Cochrane (2001). Of course, for Tables 8 to 13 , when $\widehat{M_{\mathrm{t}}^{U S}}$ and $\widetilde{M_{\mathrm{t}}^{\text {US }}}$ are used in testing, we avoid even further this possibility.

In this out of sample case, there is overwhelming evidence that return-based pricing kernels price correctly excess returns. When we consider the pricing returns directly, we are not able to reject the null in which the SDF is constructed from a factor model, but when we use the SDF constructed with AFI's methodology the result is less stark, since we reject in many situations the null hypothesis that $\mathbb{E}\left\{\widehat{M_{\mathrm{t}+1}^{\mathrm{US}}} R_{\mathrm{t}+1}^{\mathrm{i}}-1\right\} \times z_{\mathrm{t}}=0$, which for one side provides evidence in favor of the high power of the tests being used and for the other side seems to indicate that this estimated SDF does not display enough variance.

Overall, the evidence in Tables 2 to 13 suggest that return-based pricing kernels do price return differentials reasonably well out-of-sample in most cases. The equity premium is correctly priced in all but the case in which $\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}$ was employed in testing. Regarding the FPP there is overwhelming evidence that the SDF's account for the forward premium. It is important to stress that this evidence is robust across SDF estimates, and more importantly, across instruments.

It is particularly important to mention the good performance of both techniques in both empirical exercises when the ratio ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ is used as an instrument in testing (18). Because we cannot reject that (14) and (18) are true, and

$$
\begin{aligned}
\operatorname{cov}\left(M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}}\right. & \left.\frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}}, \frac{\mathrm{t} F_{\mathrm{t}+1}}{S_{\mathrm{t}}}\right)= \\
& \mathbb{E}\left\{M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}} \frac{\mathrm{t} F_{\mathrm{t}+1}}{S_{\mathrm{t}}}\right\}- \\
& \mathbb{E}\left\{M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}}\right\} \mathbb{E}\left\{\frac{\mathrm{t} F_{\mathrm{t}+1}}{S_{\mathrm{t}}}\right\},
\end{aligned}
$$

we are forced to conclude that

$$
\begin{equation*}
\operatorname{cov}\left(M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}}, \frac{\mathrm{t} F_{\mathrm{t}+1}}{S_{\mathrm{t}}}\right)=0 . \tag{30}
\end{equation*}
$$

Hence, ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ has no predictive power for

$$
M_{\mathrm{t}+1} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)}{P_{\mathrm{t}+1}} \frac{\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}}{S_{\mathrm{t}}}
$$

This is an evidence that although the excess return on uncovered over covered trading with foreign bonds is very predictable, we cannot reject that the risk "adjusted" excess return is unpredictable. This result is specially important, since when this ratio is used as instrument in some consumption-based models, the overidentifying restrictions use to be reject, maybe due to its predictability power.

[^7]Our results highlight the importance of $M_{\mathrm{t}+1}$ in generating (30). To account for forward premium from the U.S. perspective, all that was needed was an appropriate pricing kernel. Since we did not provide the model for the SDF we cannot claim to have advance in solving either puzzle, but simply in relating them.

### 4.4. Discussion

The purpose in this paper is to narrow the association between the EPP and the FPP, raising the possibility that these two problems are indeed symptoms of the same illness. The illness we are considering here is the use of a consumption-based pricing kernel, consistent with a simple representative consumer framework. We approach this question in an indirect way. Previous research has shown that, when consumption-based pricing kernels are used, they cannot price correctly neither the equity premium nor the forward premium, generating the EPP and the FPP respectively. Here, we have shown that, when return-based pricing kernels are used, we are able to price very well the forward premium and reasonably well the equity premium, i.e., under return-based pricing kernels there is no FPP and little evidence of the EPP. We therefore think that it is reasonable to pose the following question: Are the forward premium and the equity premium two symptoms of the same illness? Considering the information we have so far, our answer would be on the affirmative.

Although we believe that we have now the elements to answer the question posed in the title of this paper, a different question we may ask is: is the forward premium a reward for risk taking? If we take the covariance with $M_{\mathrm{t}+1}$ as the relevant measure of risk, then our answer is yes. However, this is not without controversy. Citing Engle (1996), once again (now, p. 162) "If the [CAPM] model were found to provide a good description of excess returns in foreign exchange markets, there would be some ambiguity about whether these predicted excess returns actually represent premiums." It, thus, seems that we have to settle for the less ambitious task of answering the question in the title and waiting for the moment we finally write our successful asset pricing model to answer this new question. ${ }^{18}$

There are some recent relative success histories in writing models that can account for the EPP, although none that can claim to be entirely from first principles. As of this moment we have developed a deep understanding of the problems associated with our consumption models. Our findings suggest that FPP is simply another consequence of the inability of traditional consumption-based pricing kernels to explain differences in returns, or risk premium. ${ }^{19}$ We believe that there is no reason for the literature on the EPP and on the FPP to evolve separately as they did, since economic theory was not rejected for different reasons specific to each problem, but because of a common culprit: the use of a traditional consumption-based pricing kernel, based on a simple representative consumer framework.

[^8]
## 5. Conclusions

In this paper we ask whether the EPP and the FPP are but two manifestations of the same problem: the failure of our consumption asset pricing models to account for risk premia in various markets. We do so by exploring panels of asset returns to extract consistent estimates of realized values for the pricing kernel and checking whether the pricing kernels that account for the equity premium will also accounts for the forward premium.

Our findings strongly suggest that the answer to our question is yes, the two puzzles are but two symptoms of the same illness. Finding a model that does account for either puzzle is bound to double its prize by accounting for the other as well. Our purpose in this paper is rather modest and we do not claim to have advanced (or even have tried to advance) an explanation to either puzzle, but simply to have provided strong evidence toward the fact that the two are closely related.

Recent work by Lustig and Verdelhan (2006) claims to have found evidence that the forward premium is simply a price for risk. They estimate the coefficient of relative risk aversion to fit an Euler equation under more general preferences than those used in Mark (1985) and Hodrick (1987) using the return on eight different portfolios of foreign currencies. Although they estimate a parameter of risk aversion above 100, one of their findings is in line with our in sample exercises since, they report (p.36) that the inclusion of data on American equity does not change significantly the estimated parameters. As of this moment we do not claim to have such explanation, although, if we take the covariance with the pricing kernel to be the price of risk, our explanation is aligned with theirs. More to the point, da Costa and Matos (2006) estimate and test the capacity of various consumption models to account for the FPP and the EPP and find that the estimated parameters take very similar values for all models, suggesting that the models perform very similarly when both markets are analyzed.

## References

[1] Araujo, F., Issler, J.V., and Fernandes, M. (2005), "Estimating the stochastic discount factor without a utility function," Working Paper \# 583, Graduate School of Economics, Getulio Vargas Foundation, downloadable as http://epge.fgv.br/portal/arquivo/1817.pdf
[2] Attanasio, O. and Browning, M. (1995) "Consumption over the life cycle and over the business cycle," American Economic Review, vol. 85(5), pp. 1118-1137, 1995.
[3] Attanasio, O. and Weber, G. (1995), "Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey," Journal of Political Economy, vol. 103(6), pp. 1121-1157, 1995.
[4] Backus, D., Foresi, S. and Telmer, C. (1995), "Interpreting the forward premium anomaly," The Canadian Journal of Economics 28, 108-119.
[5] Bai, J. (2005), "Panel Data Models with Interactive Fixed Effects," Working Paper: New York University.
[6] Bailie, R. T. and Bollerslev, T. (2000), "The forward premium anomaly is not as bad as you think," Journal of International Money and Finance 19, 471-488.
[7] Bekaert, G. (1996), "The time variation of risk and return in foreign exchange markets: a general equilibrium perspective," Review of Financial Studies 2, 427-470.
[8] Bekaert G. and Hodrick, R. J. (1993), "On biases in the measurement of foreign exchange risk premiums," Journal of International Money and Finance 12, 115-138.
[9] Blundell, R. Browning, M. and Meghir, C. (1994), "Consumer demand and the life-cycle allocation of household expenditures", Review of Economic Studies 61, 57-80.
[10] Bollerslev, T., Engle, R.F. and Wooldridge, J. (1988), "A Capital asset pricing model with time varying covariances", Journal of Political Economy 96, 116-131.
[11] Brandt, M. W., Santa-Clara, P. and Cochrane, J. H. (2006), "International risk sharing is better than you think, or exchange rates are too smooth," Journal of Monetary Economics 53, 671-698.
[12] Breeden, D. (1979), "An intertemporal asset pricing model with stochastic consumption and investment opportunities", Journal of Financial Economics 7, 265-296.
[13] Browning, M., Deaton, A. and Irish, M. (1985), "A profitable approach to labor supply and commodity demands over the life-cycle," Econometrica, Vol. 53, No. 3, pp. 503-544.
[14] Burnside, C. , Eichenbaum, M. Kleshchlski, I. and Rebelo, S. (2006), "The returns to currency speculation", NBER Working paper $\mathrm{n}^{\circ} 12489$.
[15] Campbell, J. Y. (1987), "Stock returns and the term structure", Journal of Financial Economics 18, 373-399.
[16] Campbell, J. Y. and Cochrane, J. H. (1999), "By force of habit: a consumption-based explanation of aggregate market behavior." Journal of Political Economy 107, 205-251.
[17] Campbell, J. Y. and Deaton, A. (1989), "Why is consumption so smooth?", Review of Economic Studies 56, 357-373.
[18] Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997), "The econometrics of financial markets", Princeton University Press.
[19] Chamberlain, Gary, and Rothschild, Michael, (1983). "Arbitrage, Factor Structure, and MeanVariance Analysis on Large Asset Markets," Econometrica, vol. 51(5), pp. 1281-1304.
[20] Cochrane, J.H. (2001), "Asset Pricing", Princeton Univerity Press.
[21] Connor, G., and R. Korajzcyk (1986), "Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis," Journal of Financial Economics, 15, 373-394.
[22] Connor, G., and R. Korajzcyk (1988), "Risk and return in an equilibrium APT: application of a new test methodology," Journal of Financial Economics, 21, 255-290.
[23] Connor, G. and Korajczyk, R. (1993), "A test for the number of factors in am approximate factor structure", Journal of Finance 48, 1263-1291.
[24] Constantinides, G. M. and Duffie, D. (1996), "Asset pricing with heterogeneous consumers," Journal of Political Economy 104, 219-240.
[25] da Costa, C. E. E. and Matos, P. R. F. (2006), "Do equity and foreign currency risk premiums display common patterns?", Mimeo. Getulio Vargas Foundation.
[26] Cumby, R. E., (1988) "Is it risk? Explaining deviations from uncovered interest parity", Journal of Monetary Economics 22, 279-299.
[27] Deaton, A. (1985), "Panel data from time-series of cross-sections," Journal of Econometrics, vol. 30(1/2), pp. 109-26.
[28] Engel, C. (1996), "The forward discount anomaly and the risk premium: a survey of recent evidence", Journal of Empirical Finance 3, 123-192.
[29] Engle, R.F., T. Ito, and W. L. Lin, (1990) "Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market," Econometrica 58 (3), 525-542.
[30] Engle, R. and Marcucci, J. (2005) "Common features: an overview of theory and applications", Annals Issue of the Journal of Econometrics on "Common Features", forthcoming.
[31] Fama, E. F. (1975), "Short-term interest rates as predictors of inflation", American Economic Review, 65, 269-282.
[32] Fama, E. F. (1984), "Term premiums in bond returns", Journal of Financial Economics, 13, 529-546.
[33] Fama, E. F. and Farber, A. (1979), "Money, bonds and foreign exchange", American Economic Review 69, 269-282.
[34] Frankel, Jeffrey A. (1979), "The diversifiability of exchange risk", Journal of International Economics 9, 379-393.
[35] Froot, K. A. (1990), "Short rates and expected asset returns," NBER Working paper n ${ }^{0} 3247$.
[36] Gorman, W. M. (1961), "On a class of preference fields," Metroeconomica, 13, pp. 53-56.
[37] Groen, J. J. and Balakrishnan, R. (2006),"Asset price based estimates of sterling exchange rate risk premia", Journal of International Money and Finance 25, 71-92.
[38] Hansen, L. P. (1982), "Large sample properties of generalized method of moments estimators," Econometrica 50, 1029-1054.
[39] Hansen, L. P. and Hodrick, R. J. (1983), "Risk averse speculation in the forward foreign exchange market: An econometric analysis of linear models," J.A. Frenkel, ed., Exchange rates and international macroeconomics (University of Chicago Press, Chicago, IL), 113-142.
[40] Hansen, L. P. and Jagannathan, R. (1991), "Restrictions on the intertemporal marginal rates of substitution implied by asset returns", Journal of Political Economy 99, 225-262.
[41] Hansen, L. P. and Richard, S. F. (1987), "The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models," Econometrica 55, 587-613.
[42] Hansen, L. and K. Singleton (1982) "Generalized instrumental variables estimation of nonlinear expectations models," Econometrica 50(5), pp. 1269-1286.
[43] Hansen, L. and K. Singleton. (1983), "Stochastic consumption, risk aversion and the temporal behavior of asset returns". Journal of Political Economy 91(2), 249-265.
[44] Hansen, L. and K. Singleton (1984) "Erratum of the article: 'Generalized instrumental variables estimation of nonlinear expectations models,'" Econometrica 52(1), pp. 267-268.
[45] Harrison, J. M. and Kreps, D. M. (1979), "Martingales and arbitrage in multiperiod securities markets," Journal of Economic Theory 20, 249-268.
[46] Hodrick, R. J. (1981), "International asset pricing with time-varying risk premiums," Journal of International Economics 11, 573-587.
[47] Hodrick, R. J. (1987), "The empirical evidence on the efficiency of forward and futures foreign exchange markets" (Harwood, Chur).
[48] Hodrick, R. J. (1989), "US international capital flows: perspectives from rational maximizing models," Carnegie-Rochester Conference Series on Public Policy 30, 1-29.
[49] Hodrick, R. J. and Srivastava, S. (1984), "An investigation of risk premiums and expected future spot rates," Journal of International Money and Finance 5, 5-21.
[50] Huang, R. D. (1989), "An analysis of intertemporal pricing for forward exchange contracts", Journal of Finance 44, 183-194.
[51] Kocherlakota, N. R. (1996), "The equity premium puzzle: it' s still a puzzle", Journal of Economic Literature 34(1), 42-71.
[52] Korajczyk, R. A. and Viallet, C. J. (1992), "Equity risk premia and the pricing of foreign exchange risk", Journal of International Economics 33, 199-219.
[53] Krugman, P. (1981), "Consumption preferences, asset demands, and distribution effects in international financial markets", Working paper no 651 (National Bureau of Economic Research, Cambridge, MA).
[54] Lehmann, B. and Modest, D. (1988), "The empirical foundations of the arbitrage pricing theory", Journal of Financial Economics, 21, 213-254.
[55] Lettau, M. (2001), "Idiosyncratic risk and volatility bounds, or, can models with idiosyncratic risk solve the equity premium puzzle?," Federal Reserve Bank of New York and CEPR, Staff reports n ${ }^{0} 130$.
[56] Lewis, K. K. (1990), "The behavior of Eurocurrency returns across different holding periods and regimes", Journal of Finance 45, 1211-1236.
[57] Lucas, R. E. (1978), "Asset pricing in an exchange economy," Econometrica, vol. 46, pp. 1429-1445.
[58] Lucas, R. E.
Tables and Figures
Table 1: Data summary statistics

| International quarterly data: observations from 1977:I to 2004:III |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | US\$ real <br> trading of for | eturn on the covered <br> ign government bonds | $U S \$ \text { real re }$ trading of for | urn on the uncovered <br> ign government bonds | US\$ real <br> uncovered <br> of foreig | xcess return on the er the covered trading government bonds |
| Government <br> Bonds | Sample mean (\% per year) | Sample stand. deviat. <br> (\% per year) | Sample mean (\% per year) | Sample stand. deviat. <br> (\% per year) | Sample mean (\% per year) | Sample stand. deviat. (\% per year) |
| Canadian | 2.254 | 2.209 | 2.471 | 5.441 | 0.217 | 4.922 |
| Deutsch | 2.053 | 2.638 | 2.912 | 11.623 | 0.859 | 11.962 |
| Japanese | 2.568 | 1.801 | 3.836 | 12.908 | 1.268 | 12.816 |
| Swiss | 7.705 | 15.384 | 3.065 | 16.176 | -4.640 | 19.462 |
| British | 1.756 | 2.212 | 4.695 | 10.198 | 2.940 | 10.450 |

US quarterly data: observations from 1977:I to 2004:III

| US\$ real return on 90-day T-bill |  | $U S \$$ real return on S\&P500 |  | $U S \$$ real excess return |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Sample mean <br> (\% per year) | Sample stand. deviat. <br> (\% per year) | Sample mean (\% per year) | Sample stand. deviat. (\% per year) | Sample mean (\% per year) | Sample stand. deviat. (\% per year) |
| 1.762 | 1.607 | 8.666 | 16.019 | 6.904 | 15.674 |

[^9]Figure 1: SDF's estimate using AFI' s methodology and Multifactor model.


Table 2: Forward-Premium tests for the Canadian government bonds (In sample exercise)

Table 3: Forward-Premium tests for the German government bonds (In sample exercise)

| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{\mathbf{P}_{t}\left(1+\mathrm{i}_{t}^{*}\right)}{\mathbf{P}_{t+1}} \frac{\left(\mathrm{C}_{t+1}-\mathbf{S}_{t+1}\right)}{\mathbf{S}_{t}} \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test |  |  | Conditional test |  |  |  |  |
|  |  |  | $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | $z_{\mathrm{t}}$ : (German) financial instruments |  |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{\text {t }}^{\text {C }}$ | $R_{\text {t }}^{\text {U }}$ | $R_{\text {t-1 }}^{\mathrm{C}}$ | $R_{\text {t-1 }}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
| $\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}$ | Sample mean x 100 <br> ( p -value) | $\begin{aligned} & -0.280 \\ & (0.685) \end{aligned}$ | $\begin{aligned} & -0.284 \\ & (0.685) \end{aligned}$ | $\begin{aligned} & -0.285 \\ & (0.684) \end{aligned}$ | $\begin{aligned} & -0.167 \\ & (0.692) \end{aligned}$ | $\begin{aligned} & -0.290 \\ & (0.679) \end{aligned}$ | $\begin{aligned} & -0.232 \\ & (0.741) \end{aligned}$ | $\begin{aligned} & -0.306 \\ & (0.665) \end{aligned}$ | $\begin{aligned} & -0.301 \\ & (0.672) \end{aligned}$ | $\begin{aligned} & -0.296 \\ & (0.669) \end{aligned}$ |
| $\widetilde{M_{\mathrm{t}}^{\mathrm{GL}}}$ | Sample mean x 100 ( p-value) | 0.286 $(0.681)$ | 0.291 $(0.681)$ | $\begin{gathered} 0.290 \\ (0.682) \end{gathered}$ | 0.178 $(0.675)$ | 0.287 $(0.685)$ | 0.338 $(0.634)$ | 0.277 $(0.661)$ | 0.282 $(0.695)$ | $\begin{gathered} 0.273 \\ (0.697) \end{gathered}$ |

Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).

* Reject the null hypothesis at the $5 \%$ significance level.
Table 4: Forward-Premium tests for the Japanese government bonds (In sample exercise)

| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{M_{\mathbf{t}+1} \frac{\mathbf{P}_{t}\left(1+\mathbf{i}_{t}^{*}\right)}{\mathbf{P}_{t+1}} \frac{\left(\mathbf{F}_{t+1}-\mathbf{S}_{t+1}\right)}{\mathbf{S}_{t}} \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test <br> $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | Conditional test <br> $z_{\mathrm{t}}$ : (Japanese) financial instruments |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{\text {t }}^{\text {C }}$ | $R_{\mathrm{t}}^{\mathrm{U}}$ | $R_{\mathrm{t}-1}^{\mathrm{C}}$ | $R_{\mathrm{t}-1}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
| $\widehat{M_{\mathrm{t}}^{\text {GL }}}$ | Sample mean x 100 ( p -value) | $\begin{gathered} 0.214 \\ (0.785) \end{gathered}$ | $\begin{gathered} 0.161 \\ (0.838) \end{gathered}$ | $\begin{gathered} 0.166 \\ (0.832) \end{gathered}$ | $\begin{gathered} 0.137 \\ (0.772) \end{gathered}$ | $\begin{gathered} 0.153 \\ (0.846) \end{gathered}$ | $\begin{gathered} 0.243 \\ (0.761) \end{gathered}$ | $\begin{gathered} 0.116 \\ (0.884) \end{gathered}$ | $\begin{gathered} 0.123 \\ (0.878) \end{gathered}$ | $\begin{gathered} 0.197 \\ (0.803) \end{gathered}$ |
| $\widetilde{M_{\mathrm{t}}^{\text {GL }}}$ | Sample mean x 100 <br> ( p -value) | 0.480 $(0.489)$ | 0.426 $(0.544)$ | $\begin{gathered} 0.431 \\ (0.538) \end{gathered}$ | $\begin{gathered} 0.300 \\ (0.477) \end{gathered}$ | 0.419 $(0.552)$ | $\begin{gathered} 0.513 \\ (0.467) \end{gathered}$ | $\begin{gathered} 0.385 \\ (0.587) \end{gathered}$ | $\begin{gathered} 0.404 \\ (0.572) \end{gathered}$ | $\begin{gathered} 0.466 \\ (0.506) \end{gathered}$ |

Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).

* Reject the null hypothesis at the $5 \%$ significance level.
Table 5: Forward-Premium tests for the Swiss government bonds (In sample exercise)

| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{M_{t+1} \frac{\mathbf{P}_{t}\left(1+\mathrm{i}_{t}^{*}\right)}{\mathbf{P}_{t+1}} \frac{\left(\mathrm{C}_{t} \mathrm{~F}_{t+1}-\mathbf{S}_{t+1}\right.}{\mathrm{S}_{t}} \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test |  |  | Conditional test |  |  |  |  |
|  |  |  | $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | $z_{\mathrm{t}}$ : (Swiss) financial instruments |  |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{\mathrm{t}}^{\mathrm{C}}$ | $R_{\text {t }}^{\text {U }}$ | $R_{\mathrm{t}-1}^{\mathrm{C}}$ | $R_{\mathrm{t}-1}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
|  | Sample mean x 100 | -1.525 | -1.498 | -1.500 | -0.928 | -1.773 | -1.312 | -1.596 | -1.457 | -2.051 |
|  | ( p -value) | (0.253) | (0.270) | (0.269) | (0.255) | (0.251) | (0.296) | (0.252) | (0.271) | (0.237) |
| $\widetilde{M_{\mathrm{t}}^{\mathrm{GL}}}$ | Sample mean x 100 | -0.833 | -0.789 | -0.791 | -0.505 | -1.059 | -0.624 | -0.874 | -0.724 | -1.319 |
|  | ( p -value) | (0.521) | (0.550) | (0.549) | (0.523) | (0.481) | (0.614) | (0.520) | (0.572) | (0.426) |



Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).
Table 6: Forward-Premium tests for the British government bonds (In sample exercise)

| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{M_{\mathbf{t}+1} \frac{\mathbf{P}_{t}\left(1+\mathrm{i}_{t}^{*}\right)}{\mathbf{P}_{t+1}} \frac{\left(\mathbf{F}_{t+1}-\mathbf{S}_{t+1}\right)}{\mathbf{S}_{t}} \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test <br> $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | Conditional test <br> $z_{\mathrm{t}}$ : (British) financial instruments |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{\mathrm{t}}^{\text {C }}$ | $R_{\mathrm{t}}^{\mathrm{U}}$ | $R_{\mathrm{t}-1}^{\mathrm{C}}$ | $R_{\mathrm{t}-1}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
| $\widehat{M_{\mathrm{t}}^{\mathrm{GL}}}$ | Sample mean x 100 ( p -value) | $\begin{gathered} 0.510 \\ (0.417) \end{gathered}$ | $\begin{gathered} 0.450 \\ (0.478) \end{gathered}$ | $\begin{gathered} 0.450 \\ (0.476) \end{gathered}$ | $\begin{gathered} 0.310 \\ (0.420) \end{gathered}$ | $\begin{gathered} 0.438 \\ (0.489) \end{gathered}$ | $\begin{gathered} 0.510 \\ (0.420) \end{gathered}$ | $\begin{gathered} 0.428 \\ (0.503) \end{gathered}$ | $\begin{gathered} 0.413 \\ (0.519) \end{gathered}$ | $\begin{gathered} 0.489 \\ (0.433) \end{gathered}$ |
| $\widetilde{M_{\mathrm{t}}^{\mathrm{GL}}}$ | Sample mean x 100 ( p -value) | $\begin{gathered} 0.944 \\ (0.153) \end{gathered}$ | $\begin{gathered} 0.884 \\ (0.183) \end{gathered}$ | $\begin{gathered} 0.883 \\ (0.184) \end{gathered}$ | 0.575 $(0.154)$ | 0.872 $(0.190)$ | $\begin{gathered} 0.951 \\ (0.157) \end{gathered}$ | $\begin{gathered} 0.866 \\ (0.198) \end{gathered}$ | $\begin{gathered} 0.860 \\ (0.203) \end{gathered}$ | $\begin{gathered} 0.918 \\ (0.161) \end{gathered}$ |

Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987)

* Reject the null hypothesis at the $5 \%$ significance level.
Table 7: Equity-Premium tests (In sample exercise)

| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{\left(\mathrm{i}_{t+1}^{S P}-\mathbf{i}_{t+1}^{b}\right) \mathbf{P}_{t}}{\mathbf{P}_{t+1}} \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test <br> $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | Conditional test <br> $z_{\mathrm{t}}$ : (American) financial instruments |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{t}^{\text {b }}$ | $R_{\text {t }}^{\text {S }}$ | $R_{\text {t-1 }}^{\mathrm{b}}$ | $R_{\text {t-1 }}^{\text {SP }}$ |
| $\widehat{M_{\mathrm{t}}^{\text {GL }}}$ | Sample mean x 100 <br> ( p -value) | $\begin{gathered} 1.485 \\ (0.056) \end{gathered}$ | $\begin{gathered} 1.597 \\ (0.041)^{*} \end{gathered}$ | $\begin{gathered} 1.596 \\ (0.042)^{*} \end{gathered}$ | $\begin{aligned} & 0.909 \\ & (0.054) \end{aligned}$ | $\begin{gathered} 1.598 \\ (0.042)^{*} \end{gathered}$ | $\begin{gathered} 1.640 \\ (0.040)^{*} \end{gathered}$ | $\begin{gathered} 1.593 \\ (0.044)^{*} \end{gathered}$ | $\begin{gathered} 1.629 \\ (0.040)^{*} \end{gathered}$ |
| $\widetilde{M_{\mathrm{t}}{ }^{\text {GL }}}$ | $\begin{aligned} & \text { Sample mean x } 100 \\ & (\mathrm{p} \text {-value) } \end{aligned}$ | $\begin{gathered} 0.428 \\ (0.623) \end{gathered}$ | $\begin{gathered} 0.533 \\ (0.544) \end{gathered}$ | $\begin{gathered} 0.531 \\ (0.546) \end{gathered}$ | $\begin{gathered} 0.266 \\ (0.615) \end{gathered}$ | $\begin{gathered} 0.532 \\ (0.545) \end{gathered}$ | $\begin{gathered} 0.550 \\ (0.541) \end{gathered}$ | $\begin{gathered} 0.561 \\ (0.561) \end{gathered}$ | $\begin{gathered} 0.552 \\ (0.536) \end{gathered}$ |

[^10]Table 8: Forward-Premium tests for the Canadian government bonds (Out of sample exercise)

| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} 0=\mathbb{E}\left\{M_{t+1} \frac{\mathbf{P}_{t}\left(1+\mathrm{i}_{t}^{*}\right)}{\mathbf{P}_{t+1}} \frac{\left({ }_{(t+1} \mathrm{F}_{t+1}-\mathbf{S}_{t+1}\right)}{\mathbf{S}_{t}} \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test <br> $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | Conditional test <br> $z_{\mathrm{t}}$ : (Canadian) financial instruments |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{\text {t }}^{\text {C }}$ | $R_{\mathrm{t}}^{\mathrm{U}}$ | $R_{\text {t-1 }}^{\mathrm{C}}$ | $R_{\mathrm{t}-1}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
| $\widehat{M_{\mathrm{t}}^{\text {US }}}$ | Sample mean x 100 <br> ( p -value) | $\begin{gathered} 0.010 \\ (0.978) \end{gathered}$ | $\begin{gathered} -0.050 \\ (0.989) \end{gathered}$ | $\begin{aligned} & -0.006 \\ & (0.985) \end{aligned}$ | $\begin{aligned} & -0.003 \\ & (0.989) \end{aligned}$ | $\begin{aligned} & -0.007 \\ & (0.985) \end{aligned}$ | $\begin{aligned} & -0.005 \\ & (0.990) \end{aligned}$ | $\begin{aligned} & -0.031 \\ & (0.929) \end{aligned}$ | $\begin{aligned} & -0.028 \\ & (0.938) \end{aligned}$ | $\begin{aligned} & -0.014 \\ & (0.968) \end{aligned}$ |
|  |  |  |  |  |  |  |  |  |  |  |
| $\widetilde{M_{\mathrm{t}}^{U S}}$ | Sample mean x 100 <br> ( $p$-value) | $\begin{gathered} 0.068 \\ (0.868) \end{gathered}$ | $\begin{gathered} 0.074 \\ (0.859) \end{gathered}$ | $\begin{gathered} 0.074 \\ (0.860) \end{gathered}$ | $\begin{gathered} 0.045 \\ (0.858) \end{gathered}$ | $\begin{gathered} 0.072 \\ (0.863) \end{gathered}$ | 0.076 | 0.056 | 0.061 | 0.064 |
|  |  |  |  |  |  |  | (0.858) | (0.895) | (0.885) | (0.876) |

Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).
${ }^{*}$ Reject the null hypothesis at the $5 \%$ significance level.
Table 9: Forward-Premium tests for the German government bonds (Out of sample exercise)

Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).

* Reject the null hypothesis at the $5 \%$ significance level.
Table 10: Forward-Premium tests for the Japanese government bonds (Out of sample exercise)

| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{M_{t+1} \frac{\mathbf{P}_{t}\left(1+\mathrm{i}_{t}^{*}\right)}{\mathbf{P}_{t+1}} \frac{\left(\mathrm{t}_{t} \mathrm{~F}_{t+1}-\mathbf{S}_{t+1}\right)}{\mathbf{S}_{t}} \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test <br> $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  |  |  | ditional |  |  |
|  |  |  |  |  |  |  | $z_{\mathrm{t}}$ : (Jap | financi | ruments |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{\text {t }}^{\text {C }}$ | $R_{\mathrm{t}}^{\mathrm{U}}$ | $R_{\text {t-1 }}^{\mathrm{C}}$ | $R_{\mathrm{t}-1}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
| $\widehat{M_{\mathrm{t}}^{\mathrm{US}}}$ | Sample mean x 100 ( p -value) | $\begin{gathered} 0.022 \\ (0.978) \end{gathered}$ | $\begin{gathered} 0.138 \\ (0.859) \end{gathered}$ | $\begin{gathered} 0.141 \\ (0.856) \end{gathered}$ | $\begin{gathered} 0.020 \\ (0.966) \end{gathered}$ | $\begin{gathered} 0.138 \\ (0.860) \end{gathered}$ | $\begin{gathered} 0.155 \\ (0.843) \end{gathered}$ | $\begin{gathered} 0.140 \\ (0.860) \end{gathered}$ | $\begin{gathered} 0.067 \\ (0.932) \end{gathered}$ | $\begin{gathered} 0.011 \\ (0.989) \end{gathered}$ |
| $\widetilde{M_{\mathrm{t}}{ }^{\text {US }}}$ | Sample mean $\times 100$ $(\mathrm{p}$-value $)$ | -0.075 $(0.925)$ | 0.033 $(0.967)$ | 0.036 $(0.967)$ | -0.035 $(0.947)$ | 0.029 $(0.972)$ | 0.052 $(0.949)$ | 0.031 $(0.970)$ | -0.031 $(0.970)$ | -0.090 $(0.911)$ |

Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).

* Reject the null hypothesis at the $5 \%$ significance level.
Table 11: Forward-Premium tests for the Swiss government bonds (Out of sample exercise)

|  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test <br> $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | Conditional test <br> $z_{t}$ : (Swiss) financial instruments |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\text {t }} / Y_{\text {t }}$ | $R_{\text {t }}^{\text {C }}$ | $R_{\text {t }}^{U}$ | $R_{\text {t-1 }}^{\mathrm{C}}$ | $R_{\mathrm{t}-1}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
| $\widehat{M_{\mathrm{t}}^{\text {US }}}$ | Sample mean x 100 <br> ( p -value) | $\begin{aligned} & -1.692 \\ & (0.378) \end{aligned}$ | $\begin{aligned} & -1.770 \\ & (0.363) \end{aligned}$ | $\begin{aligned} & -1.772 \\ & (0.363) \end{aligned}$ | $\begin{aligned} & -1.023 \\ & (0.382) \end{aligned}$ | $\begin{aligned} & -2.214 \\ & (0.336) \end{aligned}$ | $\begin{aligned} & -1.566 \\ & (0.370) \end{aligned}$ | $\begin{aligned} & -2.000 \\ & (0.321) \end{aligned}$ | $\begin{aligned} & -1.803 \\ & (0.330) \end{aligned}$ | $\begin{aligned} & -2.517 \\ & (0.335) \end{aligned}$ |
|  |  |  |  |  |  |  |  |  |  |  |
| $\widetilde{M_{\mathrm{t}}^{\text {US }}}$ | Sample mean x 100 <br> ( p -value) | $\begin{aligned} & -1.580 \\ & (0.442) \end{aligned}$ | $\begin{aligned} & -1.653 \\ & (0.428) \end{aligned}$ | $\begin{aligned} & -1.657 \\ & (0.427) \end{aligned}$ | -0.954 | -2.189 | -1.418 | $-1.863$ | -1.597 | -2.503 |
|  |  |  |  |  | (0.446) | (0.385) | (0.449) | (0.391) | (0.412) | (0.377) |


| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{\left[M_{\mathrm{t}+1} \frac{\mathbf{P}_{t}\left(1+\mathrm{i}_{t}^{*}\right)_{t} \mathbf{F}_{t+1}}{\mathbf{P}_{t+1} \mathbf{S}_{t}}-1\right] \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test <br> $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | Conditional test <br> $z_{\mathrm{t}}$ : (Swiss) financial instruments |  |  |  |  |
| in testing |  |  |  |  |  |  |  |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{\mathrm{t}}^{\mathrm{C}}$ | $R_{\mathrm{t}}^{\mathrm{U}}$ | $R_{\mathrm{t}-1}^{\mathrm{C}}$ | $R_{\mathrm{t}-1}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
| $\widehat{M_{\mathrm{t}}^{\mathrm{US}}}$ | $\begin{aligned} & \text { Sample mean x } 100 \\ & (\mathrm{p} \text {-value }) \end{aligned}$ | $\begin{aligned} & -1.033 \\ & (0.634) \end{aligned}$ | $\begin{aligned} & -1.052 \\ & (0.643) \end{aligned}$ | $\begin{aligned} & -1.029 \\ & (0.643) \end{aligned}$ | $\begin{aligned} & -0.620 \\ & (0.639) \end{aligned}$ | $\begin{aligned} & -0.340 \\ & (0.901) \end{aligned}$ | $\begin{aligned} & -1.327 \\ & (0.523) \end{aligned}$ | $\begin{aligned} & -0.649 \\ & (0.790) \end{aligned}$ | $\begin{aligned} & -1.192 \\ & (0.567) \end{aligned}$ | $\begin{aligned} & -0.055 \\ & (0.985) \end{aligned}$ |
| $\widetilde{M_{\mathrm{t}}^{U S}}$ | Sample mean x 100 <br> ( p -value) | $\begin{gathered} 0.854 \\ (0.853) \end{gathered}$ | $\begin{gathered} 0.998 \\ (0.832) \end{gathered}$ | $\begin{gathered} 0.985 \\ (0.834) \end{gathered}$ | $\begin{gathered} 0.534 \\ (0.848) \end{gathered}$ | $\begin{gathered} 2.237 \\ (0.663) \end{gathered}$ | $\begin{gathered} 0.195 \\ (0.967) \end{gathered}$ | $\begin{gathered} 2.082 \\ (0.673) \end{gathered}$ | $\begin{gathered} 1.155 \\ (0.802) \end{gathered}$ | $\begin{gathered} 2.364 \\ (0.642) \end{gathered}$ |

Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).
${ }^{*}$ Reject the null hypothesis at the $5 \%$ significance level.
Table 12: Forward-Premium tests for the British government bonds (Out of sample exercise)

| Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{\mathbf{P}_{t}\left(1+\mathrm{i}_{t}^{*}\right)}{\mathbf{P}_{t+1}} \frac{\left({ }_{t} \mathbf{F}_{t+1}-\mathbf{S}_{t+1}\right)}{\mathbf{S}_{t}} \times z_{\mathrm{t}}\right\}$. |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| SDF used in testing |  | Uncondit. <br> test $z_{\mathrm{t}}=1$ | Conditional test <br> $z_{\mathrm{t}}$ : Macroeconomic instruments |  |  | Conditional test <br> ritish) financial instruments |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | $C_{\mathrm{t}} / C_{\mathrm{t}-1}$ | $Y_{\mathrm{t}} / Y_{\mathrm{t}-1}$ | $C_{\mathrm{t}} / Y_{\mathrm{t}}$ | $R_{\text {t }}^{\text {C }}$ | $R_{\mathrm{t}}^{\mathrm{U}}$ | $R_{\text {t-1 }}^{\mathrm{C}}$ | $R_{\mathrm{t}-1}^{\mathrm{U}}$ | ${ }_{\mathrm{t}} F_{\mathrm{t}+1} / S_{\mathrm{t}}$ |
| $\widehat{M_{\mathrm{t}}^{\text {US }}}$ | Sample mean x 100 <br> ( $p$-value) | 0.920 | 0.924 | 0.924 | 0.562 | 0.929 | 0.948 | 0.792 | 0.763 | 0.917 |
|  |  | (0.096) | (0.102) | (0.101) | (0.094) | (0.099) | (0.099) | (0.120) | (0.143) | (0.093) |
| $\widetilde{M_{\mathrm{t}}^{U S}}$ | Sample mean x 100 ( p-value) | 1.094 | 1.102 | 1.102 | 0.669 | 1.011071 | 1.135 | 1.101 | 0.997 | 1.087 |
|  |  | (0.072) | (0.077) | (0.076) | (0.071) | (0.075) | (0.074) | (0.085) | (0.097) | (0.071) |

Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987)

* Reject the null hypothesis at the $5 \%$ significance level.
Table 13: Equity-Premium tests (Out of sample exercise)

|  | Null hypothesis in two sided "zero mean" pricing test ${ }^{1} \quad 0=\mathbb{E}\left\{M_{\mathrm{t}+1} \frac{\mathrm{i}_{t+1}^{S P}-\mathrm{i}_{t+1}^{b} \mathrm{P}_{t}}{\mathrm{P}_{t+1}} \times z_{\mathrm{t}}\right\}$. |  |
| :--- | :---: | :---: |
| SDF used | Uncondit. | Conditional test |
| in testing | test | $z_{\mathrm{t}}:$ |

## CHAPTER 2

# Do Equity and the Foreign Currency Risk Premiums Display Common Patterns? ${ }^{20}$ 


#### Abstract

In da Costa et al. (2006) we have shown how a same pricing kernel can account for the excess returns of the S\&P500 over the US short term bond and of the uncovered over the covered trading of foreign government bonds. In this paper we estimate and test the overidentifying restrictions of Euler equations associated with six different versions of the Consumption Capital Asset Pricing Model. Our main finding is that the same (however often unreasonable) values for the parameters are estimated for all models in both markets. In most cases, the rejections or otherwise of overidentifying restrictions occurs for the two markets, suggesting that success and failure stories for the equity premium repeat themselves in foreign exchange markets. Our results corroborate the findings in da Costa et al. (2006) that indicate a strong similarity between the behavior of excess returns in the two markets when modeled as risk premiums, providing empirical grounds to believe that the proposed preference-based solutions to puzzles in domestic financial markets can certainly shed light on the Forward Premium Puzzle.


JEL Code: G12;G15
Keywords: Equity Premium Puzzle, Forward Premium Puzzle, CCAPM, Habit Formation, External Timevarying Reference Consumption Level.

## 1. Introduction

Two of the most famous puzzles in financial economics are the Equity Premium Puzzle-EPP-and The Forward Premium Puzzle-FPP.

The EPP, commonly associated with the works of Hansen and Singleton (1983) and Mehra and Prescott (1985), is the failure of the traditional Consumption Capital Asset Pricing-CCAPM-to account for the excess return of stock market with respect to the risk free bond in the United States, with reasonable preference parameters. The FPP, on the other hand, relates to the difference between the forward rate and the expected future value of the spot exchange rate in a world with rational expectations and risk neutrality. ${ }^{21}$ More important for our purposes is the failure of the canonical CCAPM to generate a pattern of prices for risk that could explain the FPP being only related to the implicit risk neutrality assumption.

Despite the skepticism about risk based explanations for the FPP from part of the literature, in da Costa et al. (2006), we try to relate the puzzles. We are not able to reject that the same stochastic discount factor-SDFthat explains the equity premium also explains the forward premium. We do not use a model to investigate the puzzles. Instead we use statistical methods to extract a time series for the realized SDF and show how the moment restrictions on discounted excess returns on both equity and forward markets are not rejected. Excess returns in both markets are simply a premium for risk, if we accept the covariance with the SDF as the right measure of risk.

[^11]Because we do not spell out a full specified model it is hard to justify our calling the covariance of returns with the SDF as a risk measure. In this paper, we take the discussion of the previous paper one step further by evaluating the performance of different models in pricing excess returns for each market. Once again, our goal is not to find a model that solves all the problems. Rather, our concern is to verify if the same failures and successes attained by the CCAPM in its various forms in pricing the excess returns of equity over short term risk-free bonds will be manifest in the case of forward exchange markets.

We estimate and test for the two markets six different consumption-based asset pricing models using the Generalized Method of Moments-GMM: the canonical CCAPM, Abel's (1990) catching up with the Joneses, Abel's (1999) keeping up with the Joneses models, Epstein and Zin's (1991) recursive utility specification, Campbell and Cochrane's (1999) slow-moving habit formation model and a simplified version of Garcia's (2006) model which nests Campbell and Cochrane (1999) and Epstein and Zin (1991) for an appropriate choice of the state variables. For the equity market we test the excess return of the S\&P500 over the short-term American government bond, while for the foreign exchange markets we consider the excess return on the uncovered over the covered trading of Canadian, German, Japanese and British government bonds jointly.

As for the instruments used to generate the over-identifying restrictions that are tested in our procedure, for both puzzles, we use the same set of macroeconomic instrumental variables proven to be useful in forecasting the SDF. We also use lags of the real returns in question as the financial variables, specific for each puzzle. This procedure allows us to verify the robustness of the empirical results and the common pattern of the equity and foreign currency premiums. In order to analyze the predictability power of the forward premium pointed as a specificity of the FPP, we also use lags of this variable instead lags of the returns in question as financial instruments.

Briefly, according to our results, i) we can evidence the importance of the absolute value of the past and current aggregate per capita consumption, when used as a reference level in the utility function for both puzzles, ii) breaking the tight link between the coefficient of relative risk aversion and the elasticity of intertemporal substitution alone can not explain none of the puzzles, but iii) when besides this separation, is also considered a slow-moving external habit, then for both puzzles the agent really seems to derive utility from the level of consumption relative to this benchmark in ratio and also in difference.

Our main findings are: i) the evidence of the strong similarity between the behavior of these excess returns when modeled as risk premiums, which allows us to provide empirical grounds to believe that the proposed preference-based solutions to puzzles in domestic financial markets can certainly shed light on the FPP and ii) the prominent role played by the forward premium, relative to lags of real returns in question, as instrument in a external habit formation approach explanation of the FPP.

In particular, we show that the generalized version of the model developed by Campbell and Cochrane (1999) - with low interest rates, roughly i.i.d. consumption growth with small volatility and acceptably low values of risk aversion - is able to explain not only the high market Sharpe ratio that characterizes the EPP and but also the predictability of the foreign government bond returns.

The remainder of the paper is organized as follows. Section 2 gives an account of the literature that tries to relate the two puzzles. Subsubsection 2.2.1 discusses some characteristics of early work with consumption based model trying to emphasize the reasons why they have encountered difficulties in explaining the stylized
facts. Subsubsection 2.2.2 describes in details the features shared by most external habit formation models that address the EPP. Finally, in section 3, the empirical exercise is performed, testing all these Euler equations with GMM. Section 4 analyses the results in order to support that the equity and the currency foreign risk premiums have a similar behavior in a preference-based approach. Conclusion is the last section.

## 2. Economic Theory and risk premiums in the equity and foreign currency markets

### 2.1. Stochastic Discount Factor and Asset returns

It is now a well known fact that, given free portfolio formation, the law of one price is equivalent to the existence of a SDF through Riesz representation theorem.

Following Harrison and Kreps (1979) and Hansen and Richard (1987), we write the asset pricing equations,

$$
\begin{equation*}
1=\mathbb{E}_{\mathrm{t}}\left[M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}\right] \tag{31}
\end{equation*}
$$

and

$$
\begin{equation*}
0=\mathbb{E}_{\mathrm{t}}\left[M_{\mathrm{t}+1}\left(R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}\right)\right] . \tag{32}
\end{equation*}
$$

Here, $\mathbb{E}_{\mathrm{t}}(\cdot)$ denotes the conditional expectation given the information available at time $t, R_{\mathrm{t}+1}^{\mathrm{i}}$ and $R_{\mathrm{t}+1}^{\mathrm{j}}$ represent, respectively, the real gross return on assets $i$ and $j$ at time $t+1$ and $M_{\mathrm{t}+1}$ is the SDF , a random variable that generates unitary prices from asset returns. Under no arbitrage, we can further guarantee that there exists a strictly positive SDF that correctly prices the returns of all assets.

An immediate consequence of (32) is that excess returns are explained by the way in which returns covary with the SDF , in fact, if we take $R^{\mathrm{j}}$ to be the return on the risk-free asset, $R^{\mathrm{f}}$, equation (32) can be expanded to read

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[R_{\mathrm{t}+1}^{\mathrm{i}}\right]-R_{\mathrm{t}}^{\mathrm{f}}=-\operatorname{cov}_{\mathrm{t}}\left(R_{\mathrm{t}+1}^{\mathrm{i}}, M_{\mathrm{t}+1}\right) \tag{33}
\end{equation*}
$$

It is then, usual practice to interpret the excess expected return as a risk premium and the covariance in the right hand side of (33) as the relevant measure of risk. Using this interpretation, da Costa et al. (2006) have shown that one can use the same measure of risk to explain the equity and the forward premiums. An important question remains open, however: what does this measure of risk measure? Without a model that generates $M_{\mathrm{t}+1}$ from the primitives (preferences, endowments, technologies, etc.) of the economy the meaning of this covariance term is somewhat elusive. Not surprisingly, it is the main purpose of all research in finance economics exactly to find a model capable of generating $M_{\mathrm{t}+1}$ from the primitives of the economy. It is also the failure of most attempts to find such a model that defines most of the puzzles in finance theory.

We shall in the rest of this section present some of the models for $M_{\mathrm{t}+1}$ written in the past quarter of century and discuss their main features and drawbacks. We shall refrain from discussing models that try to account for puzzles by incorporating market incompleteness. We do so not because we think that this in not an important issue, but because following this path requires a completely different consumption data set, which we leave for later work.

### 2.2. The consumption capital asset pricing model

The single most important advance in asset pricing from an economist's perspective was the development of the consumption capital asset pricing model associated with the names of Lucas (1978) and Breeden (1979). ${ }^{22}$

Consider an economy endowed with an infinitely lived representative consumer whose preferences are representable by a von Neumann-Morgenstern utility function. It is not hard to show that the SDF is simply the growth of the representative agent's marginal utility of consumption.

Formally,-16.734lly,-28(or)1(4)1(t)2(r)1(e)-w0501977(h)-1()-33a14qnn-1(a)1(t)-347(tct)1(f)-33o-1()c th1( $\mathbb{C}() f(t h 1(a 7(v) 29(o$
consumption, $C_{\mathrm{t}},{ }^{24}$

$$
\begin{equation*}
u\left(C_{\mathrm{t}}\right)=\frac{C_{\mathrm{t}}^{1-\square}}{(1-\alpha)} \tag{37}
\end{equation*}
$$

In this case, equation (35) can be rewritten as

$$
\begin{equation*}
0=\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{-\mathrm{\square}}\left(R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}\right)\right] \tag{38}
\end{equation*}
$$

This utility function has some other important properties: it is scale-invariant and if agents in the economy have different levels of wealth but have the same power utility, then this will also be the utility function of the representative agent. Although it has been extensively used in finance literature due to its empirical, analytical and intuitive convenience, it really does not work well in practice, there being many evidences of its incapability to account for important stylized facts. The equity and the forward premium puzzles are two of the most famous and reported empirical failures of this consumption-based approach.

To give a summary account of the theoretical and empirical issues related to the EPP and the FPP, let this representative agent be an American investor who can freely trade American and foreign assets, besides having access to forward and spot exchange rate markets. ${ }^{25}$

The consumption Euler equation (38) is useful to access the EPP, if one takes assets $i$ and $j$ to be $S \& P 500$ index and the short-term American government bond.

When analyzing the FPP, the relevant assets are the covered and uncovered trade of a foreign government bond. The real returns on these foreign assets in terms of the representative investor's numeraire can be written respectively as

$$
\begin{equation*}
R_{\mathrm{t}+1}^{\mathrm{C}}=\frac{\mathrm{t} F_{\mathrm{t}+1}\left(1+i_{\mathrm{t}}^{*}\right) P_{\mathrm{t}}}{S_{\mathrm{t}} P_{\mathrm{t}+1}} \quad \text { and } \quad R_{\mathrm{t}+1}^{\mathrm{U}}=\frac{S_{\mathrm{t}+1}\left(1+i_{\mathrm{t}}^{*}\right) P_{\mathrm{t}}}{S_{\mathrm{t}} P_{\mathrm{t}+1}} \tag{39}
\end{equation*}
$$

where ${ }_{\mathrm{t}} F_{\mathrm{t}+1}$ and $S_{\mathrm{t}}$ are the forward and spot prices of foreign currency in terms of domestic currency, $P_{\mathrm{t}}$ is the dollar price level and $i_{\mathrm{t}}^{*}$ represents nominal net return on a foreign asset in terms of the foreign investor's preferences. It is, then, possible to express the consumption Euler equation of the excess returns of uncovered over covered operation with foreign bonds as

$$
\begin{equation*}
0=\mathbb{E}_{\mathrm{t}}\left\{\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{-\square} \frac{P_{\mathrm{t}}\left(1+i_{\mathrm{t}}^{*}\right)\left[\mathrm{t} F_{\mathrm{t}+1}-S_{\mathrm{t}+1}\right]}{S_{\mathrm{t}} P_{\mathrm{t}+1}}\right\} \tag{40}
\end{equation*}
$$

along the lines of Lucas' (1978) model.
Most tests reported in the literature are based on over-identifying restrictions on estimations that use Hansen's (1982) Generalized Method of Moments (GMM). With regards to the FPP, the results obtained in Mark (1985) and Hodrick (1989) report values of the coefficient of relative risk aversion, $\alpha$, above to 40 and 60, respectively, while Engel (1996) reports some estimates for $\gamma$ in excess of 100. The findings in the case of EPP are similar, in the sense that in both cases the estimated values are above 10, which is viewed as unacceptable and unreasonable, according to Krugman (1981), Mehra and Prescott (1985) and Romer (1995).

It is beyond the scope of this paper to give a full account of the large body of research produced in the area, we must however call into attention the fact that in both cases an extremely high coefficient of risk aversion is

[^12]estimated in both cases. Part of our story will be to ask whether the same values for the parameters of different models are needed to account for both puzzles. This is not the whole story, however. Even if we are willing to accept the 'unrealistic' values for the relevant parameters, the over-identifying restrictions may be rejected. We may then check whether the circumstances in which one model is accepted (or rejected) are the same for the two markets.

There is some skepticism with this regard since the risk premiums involved in these different financial markets have their specificities: the high Sharpe ratio related to the equity premium and the predictability of foreign currency excess returns based on the respective forward discount. In da Costa et al. (2006) we have shown that the same $M_{\mathrm{t}+1}$ is able to account for risk premium in both markets. We shall now investigate in detail the relative performance of different consumption based models to explain these features. Before doing so, let us briefly describe some stylized facts regarding both markets.

Stock markets commonly show familiar patterns. Observing the US quarterly data over the period 1977:1 to 2004:3, the summary statistics reported in the table 1 are in accordance with the data widely reported in the literature. The average real return on $S \& P 500$ has been $8.67 \%$ at an annual rate, while the real return on 90-day Treasury bill has been $1.76 \%$ per year. Clearly, real stock returns are much more volatile than the US Treasury bill, which is risk-free in nominal terms. Its annualized standard deviation are, respectively, $16.02 \%$ and $1.61 \% .^{26}$ We then observe an annualized Sharpe ratio for the US stock market equal to 0.44 , which implies this value as the lower bound of the standard deviation for the SDF. Certainly, a very large value for the SDF, a random variable whose mean must be close to one and whose lower bound is zero.

Since our representative agent can freely trade domestic and foreign assets and since the SDF relates all payoffs to market prices, which necessarily includes both uncovered and covered trading of foreign government bonds, should not be important, or even possible, to reconcile the characterization of the SDF provided by foreign government bond-market data with the evidence from US stock-market data?

Over this same period, the real return on covered trading of Canadian, German, Japanese and British government bonds show mean and standard deviation quite similar to the American 90-day Treasury bill one, which is reasonable, since it reflects the covered interest rate parity in a frictionless economy. Regarding the return on uncovered trading, the means are ranging from $2.47 \%$ to $4.70 \%$, while the standard deviation from $5.44 \%$ to $12.91 \%$. For these foreign bonds, we have that the Sharpe ratios for the uncovered trading of foreign bonds range from 0.04 to 0.28 .

Shiller (1982), Hansen and Jagannathan (1991) and Cochrane and Hansen (1992) relate the EPP to the volatility of the SDF, or equivalently the volatility of the intertemporal marginal rate of substitution of a representative investor required to match Sharpe ratio of the stock market. Since the higher the Sharpe ratio, the tighter the lower bound on the volatility of the SDF, when we observe these stylized facts of the American domestic and foreign government bond markets, should we consider them as an evidence that if we were able of writing down a model that consistently generated a SDF that accounts for the EPP, this model would also explain the FPP? Bearing in mind the fact that an extremely high risk aversion level in the canonical approach for both markets we ask whether there would be other similarities between these risk premiums when other preferences are considered?

[^13]The main difficulty one must face in order to answer this question is the non-existence of a widely accepted model capable of generating the observed behavior of $M_{\mathrm{t}+1}$. On the one hand, we have the poor performance of the canonical CCAPM as evidenced by the FPP and the EPP, while on the other, we have the still incipient discussion of the underlying assumptions (and empirical consequences) of more successful consumption models as Campbell and Cochrane (1999) and Garcia (2006).

Given our purposes, of simply relating the puzzles, both the successful stories and the failures are useful. The fact that the same unacceptably high risk aversion is needed to account for the two puzzles in the case of the canonical CCAPM help us relate the two puzzles quite as much as the finding of a reasonably low parameter that accounts for the two.

Finding models that produce non-acceptable values is easy. As for the successful ones a natural route is to use the external habit formation models which, in some versions, has proven to be able to account for the EPP and check whether it works for the FPP. In the next few pages we describe some of the properties of these models.

### 2.2.2. External reference level consumption-based models

The habit formation literature emerges as a natural attempt to capture some features of the consumption behavior, such as the effect of today's consumption on tomorrow's marginal utility of consumption, or in macroeconomic terms, the perception that recessions are so feared even though the recession period may not be one of the worst periods in the history. The major theoretical papers on this subject are Ryder and Heal (1973), Sundaresan (1989), and Constantinides (1990).

The main issue to be addressed in this framework is the specification of state variables that will be incorporated in the utility function of the agent, playing the role of a reference consumption level. When the state variable, $X_{\mathrm{t}}$, depends on an agent's own consumption and the agent takes it into account when choosing how much to consume, then we have a standard internal habit model, such as those in Sundaresan (1989) and also Constantinides (1990). When this habit depends on variables which are unaffected by the agent's own choice, rather depending on what others do, we are dealing with an external habit model. The literature based on this latter framework is rather extensive, including the works of Abel (1990), Abel (1999) and Campbell and Cochrane (1999), to name a few.

We will limit our analysis to the external habit models, ${ }^{27}$ for simplicity, since our intent is to explore two other common features of the habit framework when dealing with equity and foreign currency risk premiums.

When modelling habit formation the first thing we must decide is whether to use ratio or difference. ${ }^{28}$ In the first case, with constant risk aversion, the standard time-separable power utility function becomes

$$
\begin{equation*}
u\left(C_{\mathrm{t}}, X_{\mathrm{t}}\right)=\frac{1}{(1-\alpha) \varphi}\left(\frac{C_{\mathrm{t}}}{X_{\mathrm{t}}}\right)^{1-\square} X_{\mathrm{t}}^{\prime} \tag{41}
\end{equation*}
$$

where $\alpha$ is the curvature parameter for relative consumption, while $1-\varphi$ assumes this same function for the benchmark level.

[^14]In the second case, i.e., when one is considered with the difference in consumption with respect to the reference level, the same preferences yield

$$
\begin{equation*}
u\left(C_{\mathrm{t}}, X_{\mathrm{t}}\right)=\frac{1}{(1-\alpha) \varphi}\left(C_{\mathrm{t}}-X_{\mathrm{t}}\right)^{1-\square} X_{\mathrm{t}}^{\prime} \tag{42}
\end{equation*}
$$

The second feature is related to the speed with which habit reacts to aggregate consumption, where this habit can depend on current and/or lags of the reference consumption level or it can still reacts only gradually to changes in the benchmark.

In the next subsections we describe in details the modeling strategies of the consumption reference level used in this paper, proposing possible extensions of these models and establishing how the presence of this benchmark changes the respective Euler equations.

Catching-up with the Joneses First, we start following the Abel's (1990) catching up with the Joneses model, one of the most traditional and mentioned habit formation approaches. This ratio model, which was initially introduced in order to account for the high observed value for the equity premium, assumes that the habit level depends only on the first lag of aggregate consumption, denoted by $\bar{C}$.

It is easy to see this formulation as a special case of the equation (41), where $\varphi=0$ and the level habit is given by $X_{\mathrm{t}}=\bar{C}_{\mathrm{t}-1}^{\square}$, which enables us to rewrite it as

$$
\begin{equation*}
u\left(C_{\mathrm{t}}\right)=\frac{1}{(1-\alpha)}\left(\frac{C_{\mathrm{t}}}{\bar{C}_{\mathrm{t}-1}^{\square}}\right)^{1-\mathrm{\square}} \tag{43}
\end{equation*}
$$

Since, in equilibrium the relation $C_{\mathrm{t}+1} \equiv \bar{C}_{\mathrm{t}+1}$ holds, we have that the SDF can be defined as

$$
\begin{equation*}
M_{\mathrm{t}+1}^{\mathrm{CJ}}=\beta\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{-\square}\left(\frac{C_{\mathrm{t}}}{C_{\mathrm{t}-1}}\right)^{\square(\mathrm{\square}-1)} \tag{44}
\end{equation*}
$$

Note that the special case with $\kappa=0$ corresponds to the standard time-separable model, while $\kappa=1$ corresponds to the catching up with the Joneses model, where only relative consumption matters to the agent. In our estimation exercise, we follow Fuhrer (2000) in not imposing any restriction to the values of this parameter. Our intent is to figure out whether the absolute reference level matters to the representative agent and whether the agent is of the jealous or patriotic kind.

Keeping-up with the Joneses Starting with the Abel (1990) model is an interesting choice for it has many possibilities of extension. To disentangle the relative risk aversion from the elasticity of intertemporal substitution and to deal with a problem that appears when this model is used, namely, the high volatility of the risk-free rate of return, Abel (1999) assumes that the agent may take into account not only the information available to him at time $t$, but also some information available at time $t+1$, when he forms his reference consumption level, $X_{\mathrm{t}+1}$. More specifically, this framework captures the notion that the agent's benchmark level depends on current and recent levels of consumption per capita, given by the increasing function

$$
\begin{equation*}
X_{\mathrm{t}}=\bar{C}_{\mathrm{t}} \bar{C}_{\mathrm{t}-1}^{\mathrm{l}}\left(G^{\mathrm{t}}\right)^{\mathrm{\square}} \tag{45}
\end{equation*}
$$

where $0 \leq \gamma \leq 1,0 \leq \kappa \leq 1,0 \leq \eta \leq 1$ and $G \geq 1$, which allows for the possibility that the benchmark level grows simply with the passage of time. The special case with $\kappa>0$ and $\gamma=\eta=0$ corresponds to the simple
formulation of catching up with the Joneses in Abel (1990) and when $\gamma>0$ and $\kappa=\eta=0$, we have the Gali's (1994) specification of consumption externalities.

Once again, we consider $\varphi=0$, so that this formulation is a special case of equation (41) where the utility is given by

$$
\begin{equation*}
u\left(C_{\mathrm{t}}\right)=\frac{1}{(1-\alpha)}\left(\frac{C_{\mathrm{t}}}{\bar{C}_{\mathrm{t}} \bar{C}_{\mathrm{t}-1}^{\mathrm{\square}}\left(G^{\mathrm{t}}\right)^{\square}}\right)^{1-\square} \tag{46}
\end{equation*}
$$

Since, in equilibrium the relation $C_{\mathrm{t}+1} \equiv \bar{C}_{\mathrm{t}+1}$ holds, we have that the SDF now can be defined as

$$
\begin{equation*}
M_{\mathrm{t}+1}^{\mathrm{K} \mathrm{~J}}=\beta G^{\square(\mathrm{\square}-1)}\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{-\square+(\square-1)}\left(\frac{C_{\mathrm{t}}}{C_{\mathrm{t}-1}}\right)^{\square(\mathrm{\square}-1)} \tag{47}
\end{equation*}
$$

With regards the parameters involved in this Abel (1999) preference specification, there are two important issues to be mentioned. First, there are six parameters $\beta, G, \eta, \alpha, \kappa$ and $\gamma$ that affect the SDF, but only three of them are independent parameters, which implies that, relative to the power utility the current framework introduces only one additional degree of freedom. Second, although this fact allows for a variety of interpretations of preferences, our analysis will consider the Abel's (1990) estimation results.

One of the main consequences of considering a contemporaneous reference level is that it allows for disentangling risk aversion and intertemporal elasticity of substitution. We shall return to this after discussing Epstein and Zin's formulation.

## Epstein-Zin utility specification

The inclusion of extra state variables is not the only possible approach to handling the empirical failures of the canonical CCAPM. In fact, one of the important characteristics of (37) is the fact that the coefficient of relative risk aversion is the inverse of the intertemporal marginal rate of substitution. This means that, allowing for very high risk aversion to account for the equity premium implies to accept too low an intertemporal marginal rate of substitution, the consequence of which is the appearance of a risk-free puzzle whenever one is willing to accept the high coefficient of relative risk aversion that is needed to correctly price the equity premium. To avoid this trap a successful model must disentangle the two.

This is accomplished within the framework of Epstein and Zin (1991) and Weil (1989) who, building on the work of Kreps and Porteus (1978), define more general (than von-Neumann Morgenstern) preferences which: i) preserve many of the attractive features of power utility as the scale-invariance; ii) break the tight link between the coefficient of risk aversion and the elasticity of intertemporal substitution as in Abel (1999), and; iii) provide an amplified permanent component of the SDF, satisfying the Alvarez and Jermann's (2002) criticism of the lower bound of the size of this component term. ${ }^{29}$ The key of this current approach is the link between the reference level and the return on the market portfolio.

The Epstein-Zin objective function can be written as

$$
\begin{equation*}
u\left(C_{\mathrm{t}}, \mathbb{E}_{\mathrm{t}}\left(U_{\mathrm{t}+1}\right)\right)=\left[(1-\beta) C_{\mathrm{t}}^{\frac{1-\alpha}{\vartheta}}+\beta\left(\mathbb{E}_{\mathrm{t}}\left(U_{\mathrm{t}+1}^{1-\square}\right)\right)^{\frac{1}{\vartheta}}\right]^{\frac{\vartheta}{1-\alpha}} \tag{48}
\end{equation*}
$$

where $\vartheta=(1-\alpha) /(1-\rho)$ and $1 / \rho$ is the elasticity of intertemporal substitution. In the special case with $\rho=\alpha$ one can recover easily the time-separable power utility.

[^15]One important fact to bear in mind is that the Epstein-Zin non-expected recursive utility specification is observationally equivalent to the ratio habit formation model given by equation (41) if one assumes that

$$
\begin{equation*}
\log \left(\frac{X_{\mathrm{t}+1}}{X_{\mathrm{t}}}\right)=\frac{1}{1-\varphi} \log \left(R_{\mathrm{M} ; \mathrm{t}+1}\right)+\omega \tag{49}
\end{equation*}
$$

where $\omega$ is a constant and $R_{\mathrm{M} ; \mathrm{t}+1}$ is the US real gross return on the market portfolio. The proof of this equivalence as well as the intuition of this assumption can be see in details in Garcia (2006).

Using this equivalence or the assumption that the investor has no labor income and lives entirely off financial wealth, one is able to show that the SDF can now be written as

$$
\begin{equation*}
M_{\mathrm{t}+1}^{\mathrm{EZ}}=\beta\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{\square(\mathrm{\square}-1)=(1-\mathrm{\square})} R_{\mathrm{M} ; \mathrm{t}+1}^{(\mathrm{\square}-\mathrm{\square})=(1-\mathrm{\square})} \tag{50}
\end{equation*}
$$

## Campbell and Cochrane habit formation

Several papers have followed Abel $(1990,1999)$ in proposing ratio models where the habit depends on the current or on some lags of aggregate consumption. These models however fail to account for some other "recent" stylized facts of the stock market, as the predictable variation in stock returns, pointed as an important source of stock market volatility.

An alternative and promising route to explain this variation predictability is to assume that the price of risk changes through time. Campbell and Cochrane (1999) built a model where the habit level responds only gradually to changes in consumption, giving emphasis on the role played by a time-varying risk aversion. This model makes the volatility of the SDF vary in a stochastic fashion with the business cycle pattern allowing for a understanding of stock market volatility and the high equity premium.

Campbell and Cochrane (1999) assume that the power utility of the representative agent is a function of the difference between his own level of consumption and a slow-moving habit, or a time-varying subsistence level. Consequently, as consumption declines toward the habit in a business cycle trough, the curvature of the utility function rises, so risky asset prices fall and expected returns rise. The key aspects responsible for the empirical success of this approach will become more clear as we describe the main features of the model.

First, the utility function is a special case of the utility given by (42) with $\varphi=0$.
Second, let $S_{\mathrm{t}} \equiv\left(C_{\mathrm{t}}-X_{\mathrm{t}}\right) / C_{\mathrm{t}}$ denote the surplus consumption ratio. This variable captures the relation between consumption and the habit level, so the closer its value is to one, the better is the state. Note that the curvature of the utility function, $\eta_{\mathrm{t}}$, is now time-varying and it can be related to the surplus consumption ratio by $\eta_{\mathrm{t}} \equiv \alpha / S_{\mathrm{t}}$, which implies a higher local curvature in the worst states. Since the model adopts an external habit specification in which habit is determined by the history of aggregate consumption, define $\bar{S}_{\mathrm{t}} \equiv$ $\left(\bar{C}_{\mathrm{t}}-X_{\mathrm{t}}\right) / \bar{C}_{\mathrm{t}}$. Once again, note that in equilibrium the relations $C_{\mathrm{t}+1} \equiv \bar{C}_{\mathrm{t}+1}$ and $S_{\mathrm{t}+1} \equiv \bar{S}_{\mathrm{t}+1}$ hold, which enables us to rewrite the SDF as

$$
\begin{equation*}
M_{\mathrm{t}+1}^{\mathrm{cC}}=\beta\left(\frac{C_{\mathrm{t}+1} S_{\mathrm{t}+1}}{C_{\mathrm{t}} S_{\mathrm{t}}}\right)^{-\mathrm{\square}} \tag{51}
\end{equation*}
$$

Third, with regards the consumption growth, this process is modeled as an independently and identically distributed (i.i.d.) lognormal process, ${ }^{30}$

$$
\begin{equation*}
c_{\mathbf{t}+1}-c_{\mathbf{t}}=g+\nu_{\mathbf{t}+1}, \quad \nu_{\mathbf{t}+1} \sim i i d N\left(0, \sigma^{2}\right) \tag{52}
\end{equation*}
$$

[^16]where the variables in small letters are the logs of the variables in capital letters.
Finally, in order to specify how each individual's habit $X_{\mathrm{t}}$ responds to the history of aggregate consumption $\bar{C}$, let the log surplus consumption ratio $\bar{s}_{\mathrm{t}}=\ln \left(\bar{S}_{\mathrm{t}}\right)$ evolves as a heteroskedasticity AR $(1)$ process,
\[

$$
\begin{equation*}
\bar{s}_{\mathrm{t}}=(1-\phi) \bar{s}+\phi \bar{s}_{\mathrm{t}}+\lambda\left(\bar{s}_{\mathrm{t}}\right)\left(\bar{c}_{\mathrm{t}+1}-\bar{c}_{\mathrm{t}}-g\right) \tag{53}
\end{equation*}
$$

\]

where $\phi, \bar{s}$ and $g$ are the parameters that correspond respectively to the persistence coefficient, the surplus consumption ratio steady state and the mean of the log consumption growth. ${ }^{31}$ The $\lambda\left(s_{\mathrm{t}}\right)$ term, labeled sensitive function, is specified following

$$
\begin{align*}
& \lambda\left(s_{\mathrm{t}}\right)= \begin{cases}\frac{1}{\overline{\mathrm{~s}}} \sqrt{1-2\left(s_{\mathrm{t}}-\bar{s}\right)-1}, & s_{\mathrm{t}} \leq s_{\max } \\
0, & s_{\mathrm{t}} \geq s_{\max }\end{cases}  \tag{54}\\
& \text { where } \quad s_{\max } \equiv \bar{s}+\frac{1}{2}(1-\bar{S}) \text { and } \bar{S}=\sigma \sqrt{\frac{\square}{1-\square}}
\end{align*}
$$

in order to produce a constant risk-free rate and to restrict habit behavior to keep the specification close to the traditional and sensible notions on habit. One of the main purposes of the Campbell and Cochrane (1999) model, which has influenced the choice of the sensitivity function and of the parameters, is to guarantee the constancy of the risk-free rate.

## Generalized Campbell and Cochrane model

Reference models such as that of Cochrane and Campbell (1999) assume that agents derive utility from the relationship between their private consumption and the reference level. They do not allow for the reference level to directly affect utility. ${ }^{32}$ In a recent paper, Garcia (2006) have found evidence that the reference level plays an independent role in the agent's utility function. We follow them in proposing an extension of Campbell and Cochrane (1999) model in which the agent derives utility from its own level of consumption relative to the reference level both in the ratio and in difference forms. This more general framework is a special case of the setting proposed by Garcia (2006) with $\varphi=\alpha-1$.

In our generalized setting, the utility function and the SDF can be written respectively as

$$
\begin{equation*}
u\left(C_{\mathrm{t}}\right)=\frac{1}{(1-\alpha)}\left(\frac{C_{\mathrm{t}}-X_{\mathrm{t}}}{X_{\mathrm{t}}}\right)^{1-\mathrm{\square}} \tag{55}
\end{equation*}
$$

and

$$
\begin{equation*}
M_{\mathrm{t}+1}^{\mathrm{CCG}}=\beta\left(\frac{C_{\mathrm{t}+1} S_{\mathrm{t}+1}}{C_{\mathrm{t}} S_{\mathrm{t}}}\right)^{-\mathrm{\square}}\left(\frac{X_{\mathrm{t}+1}}{X_{\mathrm{t}}}\right)^{(\mathrm{\square}-1)} \tag{56}
\end{equation*}
$$

## 3. Empirical application

In this section, we estimate the Euler equations derived from the six consumption-based models described in section 2. Our intent is to either validate or better understand the reasons for the empirical failure of each preference specification, first for the excess return on the S\&P500 over the short-term American government

[^17]bond and, second, for the excess returns on the uncovered over the covered trading of Canadian, German, Japanese and British government bonds jointly.

We first describe the details and sources of the data that is used. Next, we discuss the estimation procedure in light of the identification issues surrounding the preference parameters. We shall also try and justify the criteria used in the choice of the instruments.

### 3.1. Data and instruments

In principle, whenever econometric or statistical tests are performed, it is preferable to employ a long data set either in the time-series or in the cross-sectional dimension. There are some obvious limitations, especially when the FPP is concerned, the main one being due to the forward rate series, since a sufficiently large timeseries is hardly available ${ }^{33}$. Regarding the EPP these limitations are less severe. In order to have a common sample we covered the period starting in 1977:1 and ending in 2004:3, with quarterly frequency. We collected foreign exchange data for the following countries: Canada, Germany, Japan, U.K. and the U.S.

When testing the EPP or the FPP we need first to compute excess returns $R_{t}^{i}-R_{t}^{\mathrm{j}}$. On that regard, our data set is composed of the following. To study the EPP we used the US $\$$ real returns on the S\&P500 and on 90-day T-Bill. To analyze the FPP, we used the US\$ real returns on short-term British, Canadian, German and Japanese government bonds, where both spot and forward exchange-rate data were used to transform returns denominated in foreign currency into US\$. The forward rate series were extracted from the Chicago Mercantile of Exchange database, while the spot rate series were extracted from Bank of England database.

As for macroeconomic variables, we have used the consumer price index of services and nondurable goods in the US in order to compute the $U S \$$ real returns besides the seasonally adjusted US per capita consumption of services and nondurable goods and US per capita gross domestic product. All of these variables were extracted from FED's FRED database.

The final ingredient for testing these two puzzles consists in the additional variables used in orthogonality tests. In choosing these instrumental variables, we follow Mark (1985), since we believe that the instruments he used, which are macroeconomic and financial variables observed at time $t$ by all agents, are among the most important variables used in forecasting, respectively, the SDF and the returns in question. Consequently, this set may be capable of characterizing the conditional distribution of the discounted excess returns.

For the equity premium, we use the instrument sets $\mathrm{IS}_{1}$ and $\mathrm{IS}_{2}$, given respectively by
$\mathrm{IS}_{1}$ : Two lags of the real $U S \$$ returns on $S \& P 500$ and on the 90 -day T-Bill
and three lags of real consumption and real GDP growth rates.
$\mathrm{IS}_{2}$ : Three lags of the real $U S \$$ returns on $S \& P 500$ and on the 90 -day T-Bill
and three lags of real consumption and real GDP growth rates.

[^18]For the foreign currency premiums, we use the instrument sets $\mathrm{IS}_{3}$ and $\mathrm{IS}_{4}$, given respectively by
IS $_{3}$ : Two lags of the real $U S \$$ returns on the 90 -day T-Bill and on uncovered trading of short-term British, Canadian, German, and Japanese government bonds and three lags of real consumption and real GDP growth rates.

IS 4 : Three lags of the real $U S \$$ returns on the 90 -day T-Bill and on uncovered trading of short-term British, Canadian, German, and Japanese government bonds and three lags of real consumption and real GDP growth rates.

This procedure allows us to verify the robustness of the empirical results and the common pattern of the equity and foreign currency premiums. ${ }^{34}$

In order to analyze the predictability power of the forward premium pointed as a specificity of the FPP, we also use the instrument sets $\mathrm{IS}_{5}$ and $\mathrm{IS}_{6}$, given respectively by

IS $_{5}$ : The current and the first lag of the forward premiums on the currencies in question
and three lags of real consumption and real GDP growth rates.
$\mathrm{IS}_{6}$ : The current, the first and the second lags of the forward premiums on the currencies in question
and three lags of real consumption and real GDP growth rates.

Regarding to the instruments, when we alter the set, adding lags of the financial variables our intent is to evidence the robustness of the results, while when we change all the financial variables, which happens only with the analysis os the FPP, we intend to analyze the relative predictability power of the financial variables in question.

The typical procedure adopted to validate or to better understand the reasons of the empirical failure of each preference specification for the excess returns in question consists in verifying if the model is supported by the data, in the sense of rejecting (or not) the over-identifying restriction and of analyzing the values of the parameters estimated in the light of the literature besides their significance at the $5 \%$ level. The results for this first model are reported in table 2.

There is a complete similarity between the methodologies used to estimate the canonical Euler equations and the Euler equations derived from the catching up with the Joneses model, given by

$$
\begin{equation*}
E_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{-\mathrm{\square}}\left(\frac{C_{\mathrm{t}}}{C_{\mathrm{t}-1}}\right)^{\square(\mathrm{\square}-1)}\left(R_{\mathrm{t}+1}^{\mathrm{SP}}-R_{\mathrm{t}+1}^{\mathrm{Tb}}\right)\right]=0 \tag{59}
\end{equation*}
$$

and

$$
\begin{equation*}
E_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{-\square}\left(\frac{C_{\mathrm{t}}}{C_{\mathrm{t}-1}}\right)^{\square(\mathrm{\square}-1)}\left(\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{U}}-\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{C}}\right)\right]=\overrightarrow{0}, \tag{60}
\end{equation*}
$$

and from the keeping up with the Joneses model, given by

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{\mathrm{A}}\left(\frac{C_{\mathrm{t}}}{C_{\mathrm{t}-1}}\right)^{\square}\left(R_{\mathrm{t}+1}^{\mathrm{SP}}-R_{\mathrm{t}+1}^{\mathrm{Tb}}\right)\right]=0 \tag{61}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{\mathrm{A}}\left(\frac{C_{\mathrm{t}}}{C_{\mathrm{t}-1}}\right)^{\square}\left(\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{U}}-\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{C}}\right)\right]=\overrightarrow{0} . \tag{62}
\end{equation*}
$$

The results for these models are reported respectively in tables 3 and 4 .
Now, consider the Euler equations derived from the Epstein and Zin utility specification, which are given by

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{\square(\mathrm{D}-1)=(1-\mathrm{\square})} R_{\mathrm{M} ; \mathrm{t}+1}^{(\mathrm{\square}-\mathrm{\square})=(1-\mathrm{\square})}\left(R_{\mathrm{t}+1}^{\mathrm{SP}}-R_{\mathrm{t}+1}^{\mathrm{b}}\right)\right]=0, \tag{63}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1}}{C_{\mathrm{t}}}\right)^{\square(\mathrm{\square}-1)=(1-\mathrm{\square})} R_{\mathrm{M} ; \mathrm{t}+1}^{(\mathrm{\square}-\mathrm{\square})=(1-\mathrm{\square})}\left(\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{U}}-\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{C}}\right)\right]=\overrightarrow{0} . \tag{64}
\end{equation*}
$$

For this case, the procedure adopted is the same one already described with an additional specificity, how to obtain a time series for the external reference level. In these terms, we use the value-weighted return to all stocks listed on the NYSE and AMEX obtained from the Center for Research in Security Prices (CRSP) as a proxy for the unobservable gross return on market portfolio. The results for this model are reported in table 5.

Finally, we consider the Euler equations derived form the Campbell and Cochrane (1999) model, which are given by

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1} S_{\mathrm{t}+1}}{C_{\mathrm{t}} S_{\mathrm{t}}}\right)^{-\mathrm{\square}}\left(R_{\mathrm{t}+1}^{\mathrm{SP}}-R_{\mathrm{t}+1}^{\mathrm{Tb}}\right)\right]=0 \tag{65}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1} S_{\mathrm{t}+1}}{C_{\mathrm{t}} S_{\mathrm{t}}}\right)^{-\square}\left(\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{U}}-\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{C}}\right)\right]=\overrightarrow{0} \tag{66}
\end{equation*}
$$

and the Euler equations derived from the generalized version of this latter model, given by

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1} S_{\mathrm{t}+1}}{C_{\mathrm{t}} S_{\mathrm{t}}}\right)^{-\mathrm{\square}}\left(\frac{X_{\mathrm{t}+1}}{X_{\mathrm{t}}}\right)^{(\mathrm{\square}-1)}\left(R_{\mathrm{t}+1}^{\mathrm{SP}}-R_{\mathrm{t}+1}^{\mathrm{Tb}}\right)\right]=0 \tag{67}
\end{equation*}
$$

and

$$
\begin{equation*}
\mathbb{E}_{\mathrm{t}}\left[\left(\frac{C_{\mathrm{t}+1} S_{\mathrm{t}+1}}{C_{\mathrm{t}} S_{\mathrm{t}}}\right)^{-\square}\left(\frac{X_{\mathrm{t}+1}}{X_{\mathrm{t}}}\right)^{(\square-1)}\left(\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{U}}-\tilde{\mathbf{R}}_{\mathrm{t}+1}^{\mathrm{C}}\right)\right]=\overrightarrow{0} \tag{68}
\end{equation*}
$$

When we use the Campbell and Cochrane specification and its generalized version, our main interest remains to evidence the common pattern displayed by the equity and the foreign currency risk premiums. Once again, the procedure adopted to estimate these more complex models is the typical one already described with an additional specificity related to the non-observality of the surplus consumption ratio, $S_{\mathrm{t}}$. In order to compute the time series of this process, we need to set its initial value and the values of the parameters $\sigma, g, \phi$ and $\alpha$.

With regards to the initial value of the surplus consumption ratio, we set it at the steady state value, $s_{\mathrm{o}}=\bar{s}$.
In choosing the parameters $\sigma, g$ and $\phi$, we follow Campbell and Cochrane (1999), calibrating them. Since $\sigma$ and $g$ correspond respectively to the standard deviation and the mean of the log consumption growth, we take them to match the consumption data over the period 1977:1 to 2004:3, obtaining the values 0.005 and 0.004 . With regards the persistence parameter $\phi$, we take it to match the serial correlation of log price/ dividend ratio ratios, obtaining a value of 0.989 .

For the risk aversion coefficient $\alpha$, we follow Garcia (2006), proceeding by grid search to obtain an initial value that is close to the value obtained from the estimation of the Euler equations by GMM, where $\alpha$ varies between 0.00 and 50.00.

## 4. Results

Results reported in table 2, corresponding to the Euler equations (57) and (58) derived from the canonical consumption model evidence that this model it is able to price correctly neither the excess return on the equity over the 90 -day T-Bill nor the excess return on the uncovered over covered trading of foreign government bonds with acceptably low values of risk aversion. We do not have problems with the significance of the relative risk aversion $\alpha$ and we do not reject the over-identifying restrictions. Comparing the puzzles more closely, for both of them the inclusion of the third lag of the real returns in question provides a lower risk aversion.

This empirical failure widely reported in the literature is related with the smoothness of the consumption growth along with its low correlation with the excess returns in question which require extremely high values for the risk aversion to account for the equity and the foreign currency Sharpe ratios. ${ }^{35}$

In order to try to account for the puzzles we analyze the more complex preference specifications. Table 3 presents the findings associated with Euler equations (59) and (60) derived from the Abel (1990) specification where the reference level is a function of past aggregate consumption. Comparing both puzzles: i) the EPP and the FPP are exacerbated, in the sense that, for each instrument set used, the relative risk aversion is larger then the respective one obtained with the canonical model, $i i$ ) we have the same magnitude order of the parameter $\kappa$ and iii) contrary to the ratio models, but in accordance with Garcia (2006), we find support for the hypothesis

[^19]that the absolute value of the past aggregate consumption enters the utility function. Only for the EPP, the parameter $\kappa$ is significant at the $10 \%$ significance level but not at the $5 \%$ level.

This model is not supported by the equity market neither the foreign currency market data, since the representative agent is of a jealous sort. Even with a positive $\kappa$ - where, according to Abel (1990) it would be possible to increase the risk aversion to solve the EPP without generating the risk-free rate puzzle one should not expect to explain both puzzles with this model.

In order to deal with this problem that appears when this Abel (1990) model is used, the high volatility of the risk-free rate of return, and to allow the disentangling of the relative risk aversion from the elasticity of intertemporal substitution, we estimate the Euler equations (61) and (62) derived from the Abel (1999) model. Observing the results reported in table 4, once more the equity and the foreign currency risk premiums show common patterns.

For interpretation purposes, we analyze the results in the light of Abel's (1990) results. First, we consider that only the parameters $\kappa$ and $\gamma$ are independent. In this case, with regards the past aggregate consumption the agent is a jealous sort, while with regards the current aggregate consumption the agent is patriotic. When we consider that $\alpha$ and $\gamma$ are the independent parameters, the agent remains patriotic with regards the current aggregate consumption and both puzzles are still more exacerbated.

Since, according to our results, the models which adopt state variables as a function only of societal levels of consumption are unable to support the data with reasonable values for the parameters, we test the Epstein and Zin (1991) preference which adopts a reference level as a function of the return on the market portfolio. The results of the estimation of the Euler equations (63) and (64), reported in table 5, show clearly that this attempt to break the tight link between the coefficient of risk aversion and the elasticity of intertemporal substitution does not produce a complete explanation of none of the puzzles. In fact, even when we obtaining $\rho<\alpha$, which is a necessary condition to fit the patterns of the risk premiums, both puzzles are still more exacerbated. When we are modeling the risk premiums in these different markets we obtain close values for the elasticity of intertemporal substitution, but these values are not large enough, ranging from 0.014 to 0.022 .

Our evidence about the incapacity of the Epstein and Zin (1989) framework to explain the FPP corroborates Colacito and Croce (2005), while with regards this same incapacity to explain the EPP, parts way from Epstein and Zin (1991) and Garcia (2006). ${ }^{36}$

Even though several papers in this literature have followed Abel $(1990,1999)$ in the sense of proposing ratio models where the habit depends on the current or on some lags of aggregate consumption, this path seems to be doomed, since it cannot account for the predictable variation in stock returns which is an important source of stock market volatility.

An alternative and promising route is to assume that the price of risk changes through time. With this in mind, and intending to account for the most stock market stylized facts and to better understand asset pricing puzzles, Campbell and Cochrane (1999) built a model where the habit level responds only gradually to changes in consumption, giving emphasis on the role played by a time-varying risk aversion. This model is able to make the volatility of the SDF stochastically time-varying with the business cycle pattern and to better understand the stock market volatility.

[^20]Although this model can not be considered as a complete or even as a definite model of stock market behavior, it has been pointed as the main contender to explain it. This success is certainly due to its refined and pragmatic theoretical conception besides its excellent performance in the calibration exercise proposed by Campbell and Cochrane (1999). According to the results of this exercise, the artificial data generated from the model display the patterns found in the empirical literature, but they do not estimate the coefficient of relative risk aversion and set it equal to 2 .

Here, our strategy to evidence if the risk premiums in the equity and the foreign currency markets have a similar behavior is different. We propose a GMM estimation procedure to see if these versions of this habit formation model are supported by the data, instead of calibrating them.

The estimation and test results of the Euler equations (65) and (66) derived from the Campbell and Cochrane (1999) model are reported in Table 6. For the equity and for the foreign currency risk premiums, using the lags of the returns in question as the financial instruments, the values of $\alpha$ are acceptably low but not significant at the $10 \%$ significance level. When we test the Euler equation (66) using the current and the past levels of the forward premium as the financial instruments, the values of $\alpha$ remain acceptably low but now we do not have problems with the significance. This result gives us support to emphasize the prominent role played by the forward premium, relative to lags of real returns in question, as instrument in a external habit formation approach explanation of the FPP, which may be related to the strong predictability power of the forward premium.

Comparing our results for the EPP with the result reported in the literature, while our insignificant values of $\alpha$ range from 8.01 to 8.21, Tallarini and Zhang (2005) obtains a significant $\alpha$ equal to 6.30 .

Our last attempt to model the risk premiums in these different markets relies on an extension of the Campbell and Cochrane (1999) setting, considering the case where the agent derives utility from its own level of consumption relative to the reference level both in the ratio and in difference forms. Besides allowing for a time-varying risk-free interest rate and disentangling the relative risk aversion coefficient and the elasticity of intertemporal substitution, according to the results of the Euler equations (67) and (68) reported in table 7, we can say that this extended model - with low interest rates, roughly i.i.d. consumption growth with small volatility and acceptably low values of risk aversion - is able to explain the high market Sharpe ratio that characterizes the EPP and also the strong predictability of the foreign government bond returns which characterizes the FPP.

It is hard to compare accurately our results with others reported in this literature, since the estimation strategies used and the models tested are different. At any rate, with regards the EPP, our values of $\alpha$ range from 7.64 to 8.09 , quite larger than 0.31 obtained in Garcia (2006), where the Euler equations for the excess market portfolio return and for the risk free rate are jointly estimated. About the FPP, we could compare our results with Verdelhan (2006), where the model proposed allows for a time-varying risk-free interest rate but does not break the tight link between the relative risk aversion coefficient and the elasticity of intertemporal substitution. Our values for of $\alpha$ range from 1.68 to 8.81 , while his values range from 5.60 to 8.20.

In general terms: i) we can evidence the robustness of our estimation results for the EPP and the FPP when the instrument set is altered, except for the Euler equations derived from the Abel (1999) and the Epstein and Zin (1989) models and ii) according the Hansen's test of over-identifying restrictions, neither of the six models is rejected statistically. ${ }^{37}$

[^21]Up to this point, we believe to have provided evidence in support of the fact that, under a preference-based approach, the equity and the foreign currency risk premiums display common patterns.

## 5. Conclusion

It was never clear that solving puzzles in domestic financial markets would help explaining puzzles in foreign currency markets. In fact, traditional consumption-based models do not seem to be able to account for the high equity Sharpe ratio nor to capture the non-shared characteristic of the foreign exchange market the question ceased to be posed for a long time.

Recent research, however, has lead many in the profession to believe that is a tight association between puzzles in domestic and foreign markets. We join the recent crowd in this paper adopting the following procedure: we analyze the behavior of the risk premiums in these different markets using different models and seeking not only for successes but mostly by the relative performance of different models in the two markets.

We model the excess returns of the S\&P500 over the US short term bond and of the uncovered over the covered trading of foreign government bonds, using the canonical and some external habit formation consumption approaches. Based on estimates of these consumption Euler equations and on the test of its over-identifying restrictions using Hansen's (1982) Generalized Method of Moments (GMM), we intend to validate or to better understand the reasons of the empirical failure of each preference specification for both excess returns in question.

We present an empirical evidence of the strong similarity between the behavior of the equity and the uncovered excess returns when modeled as risk premiums, which allows us to argue that, under a preference-based approach, the equity and the foreign currency risk premiums display common patterns. We find empirical grounds to believe that the proposed preference-based solutions to puzzles in domestic financial markets can certainly shed light on the FPP.

In particular, our results offers some support to the idea that a generalized version of the slow-moving external habit formation model proposed by Campbell and Cochrane (1999) which disentangles the relative risk aversion coefficient and the elasticity of intertemporal substitution is able to explain both puzzles.

## References

[1] Abel, A B. (1990), "Asset prices under habit formation and catching up with the Joneses," American Economic Review Papers and Proceedings 80, 38-42.
[2] Abel, A B. (1999), "Risk premiums and term premiums in general equilibrium," Journal of Monetary Economics 43, 3-33.
[3] Alvarez, F. and Jermann, U. J. (2002), "Using asset prices to measure the persistence of the marginal utility of wealth," unpublished paper, University of Chicago.
[4] Breeden, D. (1979), "An intertemporal asset pricing model with stochastic consumption and investment opportunities," Journal of Financial Economics 7, 265-296.
asymptotics for some of the specifications used.
[5] Burnside, C. , Eichenbaum, M. Kleshchlski, I. and Rebelo, S. (2006), "The returns to currency speculation," NBER Working paper $n^{\circ} 12489$.
[6] Campbell, J. Y. (2000), "Asset pricing at the millennium," Harvard Institute of Economic Research Discussion paper $n^{\circ} 1897$.
[7] Campbell, J. Y. and Cochrane, J. H. (1999), "By force of habit: a consumption-based explanation of aggregate market behavior," Journal of Political Economy 107, 205-251.
[8] Cochrane, J.H. (2001), "Asset Pricing," Princeton University Press.
[9] Cochrane, J. H. and Hansen, L. P. (1992), " Asset pricing lessons for macroeconomics," in O.J. Blanchard and Fischer, eds., NBER Macroeconomics Annual 1992 (MIT Press, Cambridge, MA)
[10] Colacito, R. and Croce, M. (2005), "Risks for the long run and the real exchange rate," Mimeo.
[11] Constantinides, G. M. (1982), "Intertemporal asset pricing with heterogeneous consumers and without demand aggregation," Journal of Business 55(2), 253-267.
[12] Constantinides, G. (1990), "Habit formation: a resolution of the equity premium puzzle," Journal of Political Economy 98, 519-543.
[13] da Costa, C. E. E., Issler, J. V. and Matos, P. R. F. (2006), "The forward and the equity premium puzzles: two symptoms of the same illness?" Mimeo. Getulio Vargas Foundation
[14] Engel, C. (1996), "The forward discount anomaly and the risk premium: a survey of recent evidence," Journal of Empirical Finance 3, 123-192.
[15] Epstein, L. and Zin, S. (1989), "Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework," Econometrica 57, 937-969.
[16] Epstein, L. and Zin, S. (1991), "Substitution, risk aversion and the temporal behavior of consumption and asset returns: an empirical investigation," Journal of Political Economy 99, 263-286.
[17] Fuhrer, J. (2000), "Habit formation in consumption and its implications for monetary policy models," American Economic Review 90, 367-390.
[18] Gali, J. (1994), "Keeping up with the Joneses: consumption externalities, portfolio choice and asset prices," Journal of Money Credit and Banking 26, 1-8.
[19] Garcia, R. (2006), "A consumption CAPM with a reference level," Mimeo.
[20] Gorman, W., M. (1953), "Community preference fields," Econometrica 21, 63-80.
[21] Grossman, S. J. and Shiller, R. J. (1981), "The determinants of the variability of stock market prices," American Economic Review 71, 222-227.
[22] Hansen, L. P. (1982), "Large sample properties of generalized method of moments estimators," Econometrica 50, 1029-1054.
[23] Hansen, L. P. and Richard, S. F. (1987), "The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models," Econometrica 55, 587-613.
[24] Hansen, L. P. and Jagannathan, R. (1991), "Restrictions on the intertemporal marginal rates of substitution implied by asset returns," Journal of Pol9.9Tf凡(19)1(b1.9returns99a)TJF2nal2 $\square$
[40] Rubinstein, M. (1976), "The valuation of uncertain income streams and the pricing of options," Bell Journal of Economics 7, 407-425.
[41] Ryder, H. E. and Heal, G. M. (1973), "Optimum growth with intertemporally dependent preferences," Review of Economics Studies 40, 1-33.
[42] Samuelson, P. A. (1969), "Lifetime portfolio selection by dynamic stochastic programming," Review of Economics and Statistics 51, 239-246.
[43] Shiller, R. (1982), "Consumption, asset markets and macroeconomic fluctuations," Carnegie Mellon Conference Series on Public Policy 17, 203-238.
[44] Sundaresan, S. M. (1989), "Intertemporally dependent preferences and the volatility of consumption and wealth," Review of Financial Studies 2, 73-88.
[45] Tallarini, T. D. and Zhang, H. H. (2005), "External habit and the cyclicality of expected stock returns," Journal of Business 78, 1023-1048.
[46] Verdelhan, A. (2006), "A habit-based explanation of the exchange rate risk premium," Mimeo.
[47] Weil, P. (1989), "The equity premium puzzle and the riskfree rate puzzle," Journal of Monetary Economics 24, 401-421.
Tables

Table 1: Data summary statistics*
real consumption (of nondurable and services) growth**
$0.799 \quad 0.365$
Table 2: Testing CRRA preference ${ }^{1}$


Table 3: Testing Abel (1990) catching up with the Joneses preference ${ }^{1}$

Table 4: Testing Abel (1999) keeping up with the Joneses preference ${ }^{1}$

Notes: * Coefficient not significant at the $5 \%$ significance level. ${ }^{* *}$ Reject the overidentifying restrictions at the $5 \%$ significance level. ${ }^{1}$ Hansen's (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters.
Instrument sets:
IS $_{1}$ : Two lags of the real $U S \$$ returns on $S \& P 500$ and on the 90 -day T-Bill and three lags of real consumption and real GDP growth rates. $\mathrm{IS}_{2}$ : Three lags of the real $U S \$$ returns on $S \& P 500$ and on the 90 -day T-Bill and three lags of real consumption and real GDP growth rates.
$\mathrm{IS}_{3}$ : Two lags of the real $U S \$$ returns on the 90 -day T-Bill and on uncovered trading of short-term British, Canadian, German, and Japanese government bonds and three lags of real consumption and real GDP growth rates.
IS4: Three lags of the real $U S \$$ returns on the 90 -day T-Bill and on uncovered trading of short-term British, Canadian, German, and Japanese government bonds and three lags of real consumption and real GDP growth rates.
IS $_{5}$ : The current and the first lag of the forward premia on the currencies in question and three lags of real consumption and real GDP growth rates.
IS $_{6}$ : The current, the first and the second lags of the forward premia on the currencies in question and three lags of real consumption and real GDP growth rates.
Table 5: Testing Epstein and Zin $(1989,1991)$ preference ${ }^{1}$

Table 6: Testing Campbell and Cochrane (1990) habit formation preference ${ }^{1}$


Table 7: Testing a generalized Campbell and Cochrane (1990) habit formation preference ${ }^{1}$


## CHAPTER 3

## Modeling the Foreign Risk Premiums with Time-varying Covariances


#### Abstract

The Forward Premium Puzzle (FPP) is how the empirical observation of a negative relation between future changes in the spot exchange rates and the forward premium is known. Modeling this forward bias as a risk premium, by loglinearizing the Asset Pricing Equation and under weak assumptions on the behavior of the pricing kernel, we derive an equation that describes the movement of currency depreciations, which allows us: i) to characterize the potential bias that is present in the regressions where the FPP is observed and ii) to identify the necessary and sufficient conditions that the pricing kernel has to satisfy to account for the predictability of exchange rate movements. This equation also supports that a promising path to model the exchange rate movements would be consider time-varying conditional covariances. In this sense, our main finding is that although the omitted term "explains" the risk premium in accordance with the log-linearized asset pricing equation, it has a poor performance when of its inclusion in the conventional regression without imposing any restriction about the explanatory power of the forward premium. Despite these partial results can be considered reasonable they do not seem to account for the uncomfortable forward bias, which may be better accommodated in the forthcoming extension of this working paper.


JEL Code: F31; G12; G15
Keywords: Forward Premium Puzzle, Pricing Kernel, Principal Components Analysis, Univariate GARCH approach.

## 1. Introduction

The most famous and reported puzzle in international financial economics is the forward premium puzzle henceforth, FPP - , which relates to the difference between the forward rate and the expected future value of the exchange rate in a world with rational expectations and risk neutrality. In a risk neutral world, these rates should coincide which is in contrast with the commonly reported conditional bias of forward rates as predictors of future spot exchange rates.

Despite this fact, it is important to bear in mind that, without risk neutrality, rational expectations alone does not restrict the behavior of forward rates since, as suggested by Fama (1984): it should always be possible to include a risk premium with the right properties for reconciling the time series. The rejection of the unbiasedness only represents a true puzzle if risk may cannot explain the empirical regularities found in the data. The relevant question is whether a theoretically sound economic model is able to provide a definition of risk capable of correctly pricing the forward premium. ${ }^{38}$

However, even though one can write this successful model, which has been a surmountable task, there is a robust and uncomfortable empirical finding typical of this puzzle that remains to be accommodated: the forward premium, defined as the difference between the logarithm of the forward rate and of the current exchange rate, ${ }_{\mathrm{t}} f_{\mathrm{t}+1}-s_{\mathrm{t}}$, is too strongly (negatively) correlated with subsequent changes in the (log of the) exchange rate, $s_{\mathrm{t}+1}-s_{\mathrm{t}}$. Domestic currency is expected to appreciate when domestic nominal interest rates exceed foreign

[^22]interest rates. The consequence is that interest rates differentials "predict" differential returns, a feature that is not present in domestic financial markets.

In this working paper we revisit these counterintuitive empirical findings working directly and only with the log-linearized Asset Pricing Equation. We are able to derive an equation that describes the currency depreciation movements based on a quite general framework and that has the conventional regression used in most empirical studies related to FPP as a particular case, therefore, being useful to identify the potential bias in the conventional regression due to a problem of omitted variable. ${ }^{39}$ In particular, the issue we wish to analyse here is if this bias would be responsible for the disappointing findings reported in the literature.

More closely related to our work is Korajczyk and Viallet (1992). Applying the arbitrage pricing theory APT - to a large set of assets from many countries they test whether including the factors as the prices of risk reduces the predictive power of the forward premium. They are able to show that the forward premium has lower though not zero predictive power when one includes the factors. Their results should be take with some caution since they also show that the inclusion of factors often reduces the overall explaining power (as measured by the $R^{2}$ ) of the model. Their methodology differ from ours in several dimensions.

Here, we adopt a novel three-stage approach. In order to estimate this omitted term and to analyse its relevance in accommodating the empirical evidence, first one must know the realized values of the pricing kernel. But what would be an appropriate pricing kernel?

In a recent paper, da Costa et al.(2006) circumvent the lack of a good model for the risk premium by using the principal-component and factor analyses to extract the pricing kernel, a model-free methodology that can be traced back to the work of Ross (1976), developed further by Chamberlain and Rothschild (1983), and Connor and Korajczyk $(1986,1993) .{ }^{40}$ One of their results is that applying this methodology to a data set called global assets, which is comprised of assets traded in different countries, it is not possible to reject the hypothesis that the Stochastic Discount Factor ${ }^{41}$ - henceforth, SDF — thus constructed prices correctly all returns and excess
linearized asset pricing equation, it has a poor performance when of its inclusion in the conventional regression without imposing any restriction about the explanatory power of the forward premium.

Finally, our main theoretical finding is that we are able to draw stronger implications for the SDF characterizing not only the necessary conditions that the SDF must display if it is to account for the FPP - already showed Backus et al. (1995) who first derived this log-linearized characterization used here - but also the sufficient ones.

The remainder of the paper is organized as follows. Section 2 gives a summary account of the conventional efficient test used to evidence the FPP besides laying the foundations of an equation derived from Asset Pricing Equation useful to understand some evidences related to FPP. In section 3, we describe our strategy to analyse the relevance of the inclusion of the omitted term. In section 4, the results are analyzed. The conclusion is in the last section.

## 2. Literature Overview

### 2.1. FPP: the conventional regression and previous evidences

In this brief review we emphasize the research on the FPP that tries to explain it as a premium for accepting non-diversifiable risk.

The FPP, is how one calls the systematic violation of the "efficient-market hypothesis" for foreign exchange markets, where by "efficient-market hypothesis" we mean the proposition that the expected return to speculation in the forward foreign exchange market conditional on available information should be zero. ${ }^{42}$

Most studies report the FPP through the finding of $\hat{\alpha}_{1}<1$ when running the regression,

$$
\begin{equation*}
s_{\mathbf{t}+1}-s_{\mathrm{t}}=\alpha_{0}+\left(1-\alpha_{1}\right)\left(\mathrm{t} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+u_{\mathrm{t}+1} \tag{69}
\end{equation*}
$$

where $s_{\mathrm{t}}$ is the log of time $t$ exchange rate, $f_{\mathrm{t}+1}$ is the log of time $t$ forward exchange rate contract and $u_{\mathrm{t}+1}$ is the regression error. ${ }^{43}$ Notwithstanding the possible effect of Jensen inequality terms, testing the marketefficiency hypothesis of forward exchange rate market is equivalent to testing the null $\hat{\alpha}_{1}=0, \hat{\alpha}_{0}=0$, along with the uncorrelatedness of residuals from the estimated regression. The null is rejected in almost all studies.

As pointed out by Hansen and Hodrick (1983), however, this violation of market efficiency ought not to be viewed as evidence of market failure or some form of irrationality, since the uncovered parity need only hold exactly in a world of risk neutral agents or if the return on currency speculation is not risky. ${ }^{44}$ Nevertheless, the findings came to be called a puzzle due to, at least, one good reasons: the discrepancy from the null. Although one would not be surprised with $\hat{\alpha}_{1} \neq 0$, the fact that, $\hat{\alpha}_{1}<1$, in most studies - according to Froot (1990), the average value of $\hat{\alpha}_{1}$ is $-1,88$ for over 75 published estimates across various exchange rates and time periods - implies an expected domestic currency appreciation when domestic nominal interest rates exceed foreign interest rates. Moreover, the magnitude of the coefficients suggest a behavior of a risk premium which is hard to justify with our current models.

[^23]This forward premium anomaly has become a well established regularity and according to Bekaert and Hodrick (1993) these previous inferences are shown to be robust to the measurement error due to incorrect sampling of the data or to failure to account for bid-ask spreads.

### 2.2. Economic Theory and the FPP

Since the null hypothesis of the regression (69), often rejected, represents the equilibrium condition in a world where markets are efficient and the agents are risk neutral, have rational expectations and value returns in nominal terms, there is a literature that attempts to describe the exchange rate movements assuming the conditional joint lognormality of some specific variables and relying on less restrictive assumptions with the intent to identify the reasons of the counterintuitive empirical findings reported. Along these lines, based on a more general framework as the Asset Pricing Equation, we are able to derive an equation that describes the currency depreciation capable to support a discussion about the problems related to the empirical evidences of this predictability.

### 2.2.1. Stochastic Discount Factor and Asset returns

Following Harrison and Kreps (1979) and Hansen and Richard (1987), we write the asset pricing equations,

$$
\begin{equation*}
1=\mathbb{E}_{\mathrm{t}}\left(M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}\right) \tag{70}
\end{equation*}
$$

and

$$
\begin{equation*}
0=\mathbb{E}_{\mathrm{t}}\left[M_{\mathrm{t}+1}\left(R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}\right)\right] \tag{71}
\end{equation*}
$$

Here, $\mathbb{E}_{\mathrm{t}}(\cdot)$ denotes the conditional expectation given the information available at time $t, R_{\mathrm{t}+1}^{\mathrm{i}}$ and $R_{\mathrm{t}+1}^{\mathrm{j}}$ represent, respectively, the real gross return on assets $i$ and $j$ at time $t+1$ and $M_{\mathrm{t}+1}$ is the SDF , a random variable that generates unitary prices from asset returns.

Given free portfolio formation, the law of one price - the fact that two assets with the same payoff in all states of nature must have the same price - is equivalent to the existence of a SDF through Riesz representation theorem. The correlation with the SDF is the only measure of risk that matters for pricing assets. It is, therefore, clear that we need not rely on the strong assumptions to guarantee the validity of (70) and (71). Under no arbitrage, we can further guarantee that there exists a strictly positive SDF that correctly prices the returns of all assets.

Uniqueness of $M_{\mathrm{t}+1}$, however, is harder to come about. In general, if markets are not complete there will be a continuum of SDF's pricing all traded securities. Yet, there will still exist a unique $\mathrm{SDF}, M_{\mathrm{t}+1}^{*}$, in the span of traded assets, labeled mimicking portfolio. This SDF is the unique element of the payoff space that prices all traded securities. Any $\mathrm{SDF}, M_{\mathrm{t}+1}$, may thus be written as $M_{\mathrm{t}+1}=M_{\mathrm{t}+1}^{*}+\nu_{\mathrm{t}+1}$ for some $\nu_{\mathrm{t}+1}$ with $\mathbb{E}_{\mathrm{t}}\left[\nu_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}\right]=0 \forall i$.

This approach allows us to conveniently specify the assumptions about the economy that are implicit in the choice of the function that determines the SDF with the intent to obtain further results.

To give a summary account of the theoretical and empirical issues related to these puzzle, one must recall that the covered and uncovered trade of a foreign government bond are themselves financial assets and, in this sense, the asset pricing equation (71) can be useful to analyze the FPP. ${ }^{45}$

[^24]The real returns on these foreign assets in terms of the representative investor's numeraire can be written respectively as

$$
\begin{equation*}
R_{\mathrm{t}+1}^{\mathrm{C}}=\frac{\mathrm{t} F_{\mathrm{t}+1}\left(1+i_{\mathrm{t}}^{*}\right) P_{\mathrm{t}}}{S_{\mathrm{t}} P_{\mathrm{t}+1}} \quad \text { and } \quad R_{\mathrm{t}+1}^{\mathrm{U}}=\frac{S_{\mathrm{t}+1}\left(1+i_{\mathrm{t}}^{*}\right) P_{\mathrm{t}}}{S_{\mathrm{t}} P_{\mathrm{t}+1}} \tag{72}
\end{equation*}
$$

where ${ }_{\mathrm{t}} F_{\mathrm{t}+1}$ and $S_{\mathrm{t}}$ are the forward and spot prices of foreign currency in terms of domestic currency, $P_{\mathrm{t}}$ is the dollar price level and $i_{\mathrm{t}}^{*}$ represents nominal net return on a foreign asset in terms of the foreign investor's preferences. ${ }^{46}$

### 2.2.2. Currency depreciation theoretical framework

Now consider an economy where assumptions 1 and 3 are valid.

Assumption 1: The Asset Pricing Equation $\mathbb{E}_{\mathrm{t}}\left(M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}\right)=1$ holds;

Assumption 2: There are no arbitrage opportunities;

We have many reasons to think that asset markets can be described by these assumptions. They are mild and underlie most of the recent and fundamental insights in finance; see, e.g., Hansen and Singleton (1982, 1983, 1984), Mehra and Prescott (1985) and Mulligan (2002).

In order to derive an equation that describes the currency depreciation movements, we shall add some structure on the stochastic process for the SDF and the real returns, specifying the conditional joint distribution of $M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}$ through the next assumption. Despite its prominent role in both theoretical and empirical studies of asset pricing due to its analytical convenience..

Assumption 3: The conditional joint distribution of $M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}$ is lognormal.

Taking logs in both sides of (6) and further applying a Taylor expansion yields, for every asset $i$ in the economy, we have

$$
\begin{equation*}
r_{\mathrm{t}+1}^{\mathrm{i}}=-m_{\mathrm{t}+1}-\frac{1}{2}\left(\sigma_{\mathrm{t}}^{\mathrm{i}}\right)^{2}+\varepsilon_{\mathrm{t}+1}^{\mathrm{i}} \tag{73}
\end{equation*}
$$

where the variables in small letters are the logs of the variables in capital letters, $\varepsilon_{\mathrm{t}+1}^{\mathrm{i}}=\ln \left(M_{\mathrm{t}} R_{\mathrm{t}+1}^{\mathrm{i}}\right)-$ $\mathbb{E}_{\mathrm{t}}\left\{\ln \left(M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}\right)\right\}$ denotes the innovation in predicting $\ln \left(M_{\mathrm{t}} R_{\mathrm{t}}^{\mathrm{i}}\right),\left(\sigma_{\mathrm{t}}^{\mathrm{i}}\right)^{2} \equiv \mathbb{V}_{\mathrm{t}}\left(\ln M_{\mathrm{t}+1}+\ln R_{\mathrm{t}+1}^{\mathrm{i}}\right)$ and $\mathbb{V}_{\mathrm{t}}(\cdot)$ denotes the conditional variance given the available information at time $t$. From (73) one verifies that asset returns are decomposed in three terms. The first one is the logarithm of the $\mathrm{SDF}, m_{\mathrm{t}+1}$, which is common to all returns and is a random variable. The second one, $\left(\sigma_{\mathrm{t}}^{\mathrm{i}}\right)^{2}$, is idiosyncratic and, given past information, also deterministic, but not necessarily constant. The third one is $\varepsilon_{\mathrm{t}+1}^{\mathrm{i}}$. Idiosyncratic, as well, and unforecastable, which means that it presents no serial correlation. Hence, disregarding deterministic terms, the only source of serial correlation is $m_{\mathrm{t}}$.

Because equation (73) must hold for any asset traded in an economy where assumptions 1 to 3 are valid, including foreign government bonds, it is straightforward to show that

$$
\begin{equation*}
s_{\mathrm{t}+1}-s_{\mathrm{t}}=\frac{1}{2}\left[\left(\sigma_{\mathrm{t}}^{\mathrm{C}}\right)^{2}-\left(\sigma_{\mathrm{t}}^{\mathrm{U}}\right)^{2}\right]+\left({ }_{\mathrm{t}} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+\varepsilon_{\mathrm{t}+1}^{\mathrm{U}}-\varepsilon_{\mathrm{t}+1}^{\mathrm{C}} \tag{74}
\end{equation*}
$$

[^25]where $\sigma_{\mathrm{t}}^{\mathrm{C}} \equiv \mathbb{V}_{\mathrm{t}}\left(\ln M_{\mathrm{t}+1}+\ln R_{\mathrm{t}+1}^{\mathrm{C}}\right)$ and $\sigma_{\mathrm{t}}^{\mathrm{U}} \equiv \mathbb{V}_{\mathrm{t}}\left(\ln M_{\mathrm{t}+1}+\ln R_{\mathrm{t}+1}^{\mathrm{U}}\right)$.
Consequently,
\[

$$
\begin{equation*}
s_{\mathrm{t}+1}-s_{\mathrm{t}}=-\left[\frac{1}{2} \mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)+\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)\right]+\left(\mathrm{t}_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+\varepsilon_{\mathrm{t}+1} \tag{75}
\end{equation*}
$$

\]

where $\operatorname{cov}_{\mathrm{t}}(\cdot, \cdot)$ denotes the conditional covariance given the available information at time $t, \pi_{\mathrm{t}+1}$ is the (log of) domestic price variation and $\varepsilon_{\mathrm{t}+1}$ denotes the innovation in predicting $\ln \left(M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{U}}\right)-\ln \left(M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{C}}\right)$.

This equation (75) is useful in drawing implications for a SDF that accounts for the FPP.
Taking the conditional expectation on (75) and on (69) ${ }^{47}$, it is straightforward to see that

$$
\begin{aligned}
& \mathbb{E}_{\mathrm{t}}\left(s_{\mathrm{t}+1}-s_{\mathrm{t}}\right)=- {\left[\frac{1}{2} \mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)+\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)\right]+\left({ }_{\mathrm{t}} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right) } \\
& \mathbb{E}_{\mathrm{t}}^{*}\left(s_{\mathrm{t}+1}-s_{\mathrm{t}}\right)=\alpha_{0}+\alpha_{1}\left(\mathrm{t} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)
\end{aligned}
$$

and consequently,

$$
\begin{equation*}
\left(\mathrm{t}_{\mathrm{t}+1}-s_{\mathrm{t}}\right)\left(1-\alpha_{1}\right)=\alpha_{0}+\mathbb{E}_{\mathrm{t}}^{*}\left[\frac{1}{2} \mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)+\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)\right] \tag{76}
\end{equation*}
$$

Denoting by $\xi_{\mathrm{t}}\left(m_{\mathrm{t}+1}\right) \equiv \mathbb{E}_{\mathrm{t}}^{*}\left[\frac{1}{2} \mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)+\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)\right]$ and by $\mu_{\mathrm{i}}(x) \equiv i$ th unconditional central moment of a univariate probability function $P(x)$, we can conclude that, under lognormality, to account for the FPP, i.e. to account for negative values of the parameter $\alpha_{1}$, it is necessary and sufficient that the SDF satisfies

$$
\left\{\begin{array}{cc}
\mu_{\mathrm{i}}\left(\xi_{\mathrm{t}}\left(m_{\mathrm{t}+1}\right)\right)>\mu_{\mathrm{i}}\left({ }_{\mathrm{t}} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right), & \text { for } i \text { even }  \tag{77}\\
\left|\mu_{\mathrm{i}}\left(\xi_{\mathrm{t}}\left(m_{\mathrm{t}+1}\right)\right)\right|>\left|\mu_{\mathrm{i}}\left(\mathrm{t} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)\right| \text { and } & \\
\text { both moments with the same sign, } & \text { for } i \text { odd, }
\end{array}\right.
$$

which are in accordance with Fama (1984) necessary conditions and can be considered stronger than the usual ones described in the literature. For instance, these conditions have the necessary conditions reported in Backus et al. (1995) as an implication.

This currency depreciation equation (75) is also extremely useful in deriving empirical tests of FPP. ${ }^{48}$
First, one should note that the conventional regression (69) is a particular case of (75), where the term $\frac{1}{2} \mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)+\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)$ is assumed to be time-invariant. Based on (75), i.e. in a lognormal world, it is apparent that the disappointing empirical findings reported when regression (69) is used can be due to a problem of omitted variable which necessarily makes the estimator more inefficient and creates bias and inconsistency if the omitted term is correlated with the forward premium.

Regarding this omitted term, it is worth mentioning that the $\mathbb{V}_{\mathrm{t}}\left(s_{\mathrm{t}+1}\right)$ term, which we referred to before as the Jensen's inequality term, appears only because we defined the expected rate of depreciation in logarithms. According to Bekaert and Hodrick (1993) omitting this term from regression (69) is not responsible for finding

[^26]$\hat{\alpha}_{1}<0$. Therefore, all the action must come from the term $\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)$. This term must exhibit considerable variation if one is to accommodate the evidence regarding the forward premium. Given that asset returns have clear signs of conditional heteroskedasticity, ${ }^{49}$ time invariance for the covariance term does not seem to be a sound assumption, leading one to wonder whether this might not be the key to accounting for $\hat{\alpha}_{1}<0$.

To summarize, the conventional regression used to evidence the FPP omits a conditional covariance term creating consequently a conditional bias of forward rates as predictors of future spot exchange rates.

## 3. Empirical Application

In this section we describe in details the three-stage approach used to analyse if the conditional bias can be (or not) responsible for the disappointing findings reported in the literature.

### 3.1. Return-Based Pricing Kernel

In an important paper, Hansen and Jagganathan (1991) have lead the profession towards return-based kernels instead of consumption-based kernels, something the literature on factor and principal-component analysis in Finance concurs with - e.g., Connor and Korajzcyk (1986), Chamberlain and Rothschild (1983), Bai (2005), and Araujo, Issler and Fernandes (2005). The whole idea is to combine statistical methods with economic theory to devise pricing-kernel estimates that do not depend on preferences but solely on returns, thus being projections of SDF in the space of returns, at least to a first-order approximation.

The basic idea behind return-based pricing kernels is that asset prices (or returns) convey information about the intertemporal marginal rate of substitution in consumption. In equilibrium, all returns must have a common feature (factor), which can be removed by subtracting any two returns. This common feature is the SDF. One way to see this, is to realize that although

$$
\begin{aligned}
\mathbb{E}\left\{M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{i}}\right\} & =1, \text { for all } i=1,2, \cdots, N, \\
\mathbb{E}\left\{M_{\mathrm{t}+1}\left(R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}\right)\right\} & =0, \text { for all } i \neq j .
\end{aligned}
$$

Hence, even though $R_{\mathrm{t}+1}^{\mathrm{i}}$ is not orthogonal to $M_{\mathrm{t}+1}$ itself, if we subtract the returns of any two assets, their return differential, $R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}$, will be orthogonal to $M_{\mathrm{t}+1}$. Hence, $R_{\mathrm{t}+1}^{\mathrm{i}}$ contains the feature $M_{\mathrm{t}+1}$ but $R_{\mathrm{t}+1}^{\mathrm{i}}-R_{\mathrm{t}+1}^{\mathrm{j}}$ does not.

Because every asset return $R_{\mathrm{t}+1}^{\mathrm{i}}$ contains "a piece" of $M_{\mathrm{t}+1}$, if we combine a large number of returns, the average idiosyncratic component of returns will vanish in limit. Hence, if we choose our weights properly, we may end up with the common component of returns, i.e., the SDF. Of course, this method is asymptotic, either $N \rightarrow \infty$ or $N, T \rightarrow \infty$, and relies on weak-laws of large numbers to provide consistent estimators of the mimicking portfolio - the systematic portion of asset returns.

The second uses principal-component and factor analyses to extract common components of asset returns. It can be traced back to the work of Ross (1976), developed further by Chamberlain and Rothschild (1983), and Connor and Korajczyk (1986, 1993). Additional references are Hansen and Jagganathan (1991) and Bai (2005).

[^27]Multifactor Models Factor models summarize the systematic variation of the $N$ elements of the vector $\mathbf{R}_{\mathrm{t}}=\left(R_{\mathrm{t}}^{1}, R_{\mathrm{t}}^{2}, \ldots, R_{\mathrm{t}}^{\mathrm{N}}\right)^{\prime}$ using a reduced number of $K$ factors, $K<N$. Consider a $K$-factor model in $R_{\mathrm{t}}^{\mathrm{i}}$ :

$$
\begin{equation*}
R_{\mathrm{t}}^{\mathrm{i}}=a_{\mathrm{i}}+\sum_{\mathrm{k}=1}^{\mathrm{K}} \beta_{\mathrm{i} ; \mathrm{k}} f_{\mathrm{k} ; \mathrm{t}}+\eta_{\mathrm{it}} \tag{78}
\end{equation*}
$$

where $f_{\mathrm{k} ; \mathrm{t}}$ are zero-mean pervasive factors and, as is usual in factor analysis,

$$
\operatorname{plim} \frac{1}{N} \sum_{i=1}^{\mathrm{N}} \eta_{i ; t}=0
$$

Denote by $\Sigma_{\mathbf{r}}=\mathbb{E}\left(\mathbf{R}_{\mathrm{t}} \mathbf{R}_{\mathrm{t}}^{\prime}\right)-\mathbb{E}\left(\mathbf{R}_{\mathrm{t}}\right) \mathbb{E}\left(\mathbf{R}_{\mathrm{t}}^{\prime}\right)$ the variance-covariance matrix of returns. The first principal component of the elements of $\mathbf{R}_{\mathrm{t}}$ is a linear combination $\theta^{\prime} \mathbf{R}_{\mathrm{t}}$ with maximal variance subject to the normalization that $\theta$ has unit norm, i.e., $\theta^{\prime} \theta=1$. Subsequent principal components are identified if they are all orthogonal to the previous ones and are subject to the same normalization. The first $K$ principal components of of $\mathbf{R}_{\mathrm{t}}$ are consistent estimates of the $f_{\mathrm{k} ; \mathrm{t}}$ 's. Factor loadings can be estimated consistently by simple OLS regressions of the form (78).

Hence, principal components are simply linear combinations of returns. They are constructed to be orthogonal to each other, to be normalized to have a unit length and to deal with the problem of redundant returns, which is very common when a large number of assets is considered. They are ordered so that the first principal component explains the largest portion of the sample covariance matrix of returns, the second one explains the next largest portion, and so on.

When a multi-factor approach is used, the first step is to specify the factors by means of a statistical or theoretical method and then to consider the estimation of the model with known factors. Here, we are not particularly concerned about identifying the factors themselves. Rather, we are interested in constructing an estimate of the SDF, labeled $\widetilde{M_{\mathrm{t}}^{*}}$, by collapsing the factors into a single variable. For that, we shall rely on a purely statistical model. Given estimates of the factor model of $R_{\mathrm{t}}^{\mathrm{i}}$ in (78), one can write the respective expected-beta return expression

$$
\mathbb{E}\left(R^{\mathrm{i}}\right)=\gamma+\sum_{\mathrm{k}=1}^{\mathrm{K}} \beta_{\mathrm{i} ; \mathrm{k}} \lambda_{\mathrm{k}}, i=1,2, \ldots, N
$$

where $\lambda_{\mathrm{k}}$ is interpreted as the price of the factor $k$ risk. The fact that the zero-mean factors $f \equiv \tilde{f}-\mathbb{E}(\tilde{f})$ are such that $\tilde{f}$ are returns with unitary price allows us to measure the $\lambda$ coefficients directly by

$$
\lambda=\mathbb{E}(\tilde{f})-\gamma
$$

and consequently to estimate only $\gamma$ via a cross-sectional regression ${ }^{50}$.
Given these coefficients, can easily get an estimate of $M_{\mathrm{t}}^{*}$. The relation between $(\lambda, \gamma)$ and $(a, b)$, given by

$$
a \equiv \frac{1}{\gamma} \text { and } b \equiv-\gamma\left[\operatorname{cov}\left(f f^{\prime}\right)\right]^{-1} \lambda
$$

allows us to ensure the equivalence between this beta pricing model and the linear model for the SDF, i.e. that

[^28] $\gamma$, one set it equal to the real return of the risk free rate.
$$
\widetilde{M_{\mathrm{t}}^{*}}=a+\sum_{\mathrm{k}=1}^{\mathrm{K}} b_{\mathrm{k}} f_{\mathrm{k} ; \mathrm{t}} \text { and } \mathbb{E}\left(\widetilde{M_{\mathrm{t}}^{*}} R^{\mathrm{i}}\right)=1, i=1,2, \ldots, N
$$
hold.
The number of factors used in the empirical analysis is an important issue. We expect $K$ to be rather small, but have some flexibility for this choice. We took a pragmatic view here, increasing $K$ until the estimate of $M_{\mathrm{t}}^{*}$ changed very little due to the last increment; see for instance Lehmann and Modest (1988) and Connor and Korajczyk (1988). We also performed a robustness analysis for the results of all of our statistical tests using different estimates of $M_{\mathrm{t}}^{*}$ associated with different $K$ 's. Results changed very little around our choice of $K$.

### 3.2. Conditional heteroskedasticity models

Since we have extracted a time series for the SDF, we are able to use it in order to model exchange rate movements.

Despite the equations (74) and (75) are obviously similar, they suggest different ways to obtain the omitted term in the conventional regression. The former one suggests that a correct econometrically procedure to estimate this omitted term is choosing the best models for the $\mathbb{V}_{\mathrm{t}}\left(\ln M_{\mathrm{t}+1}+\ln R_{\mathrm{t}+1}^{\mathrm{C}}\right)$ and $\mathbb{V}_{\mathrm{t}}\left(\ln M_{\mathrm{t}+1}+\ln R_{\mathrm{t}+1}^{\mathrm{U}}\right)$ terms separately, instead of working with a multivariate GARCH process required if one takes into account the latter equation.

In a first moment, we limit our empirical exercise working only with the equation (74) due to its analytical simplicity, although we still intend to analyse the individual role played by the $\operatorname{cov}_{\mathrm{t}}\left(s_{\mathrm{t}+1}, m_{\mathrm{t}+1}-\pi_{\mathrm{t}+1}\right)$ term in the equation (75) in a forthcoming extension of this current paper.

If, for instance, we are dealing with the uncovered trading of foreign government bonds, in order to satisfy the equation

$$
m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{U}}=-\frac{1}{2} \mathbb{V}_{\mathrm{t}}\left(m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{U}}\right)+\varepsilon_{\mathrm{t}+1}^{\mathrm{U}}
$$

where the expectations of the process $m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{U}}$ depends upon its conditional variance and due to our necessity to model conditional variances, it seems relevant or even necessary to use the well-known GARCH-in-mean process. Despite this specification proposed by Engle, Lilien and Robins (1987) does not follow from any economic theory, it is well-known that it provides a good approximation to the heteroskedasticity typically found in financial time-series data, since it extends the family of Autoregressive Conditional Heteroskedasticity models introduced by Engle (1982) allowing the conditional variance to affect the mean. The idea is that as the degree of uncertainty in asset returns varies over time, the consumption required by risk averse economic agents for holding these assets, must also be varying.

A problem we have to face in order to in order to model th variance of each $m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{i}}$ process is that the literature on conditional heteroskedasticity models is so extensive, ${ }^{51}$ so does the amount of such models proposed, which in some sense obligates us to choose a representative set of models to be compared.

Bearing in mind that the most widely used volatility predicting models for financial and exchange rate series are the parsimonious ones, ${ }^{52}$ to obtain the correct conditional variance model for each process, we compare some univariate parsimonious heteroskedastic models, using both the Akaike and the Schwarz criteria information.

[^29]First we consider only the ARCH and its generalized version, i.e. the GARCH model. But, since these specifications do not allow for an asymmetric effect of the arrival of good and bad news in the market on the second moment of volatility, a well-known phenomenum in financial modeling, we extend our set of models, testing two asymmetric specifications: the exponential ARCH, introduced by Nelson (1991) and the Threshold ARCH introduced independently by Zakoïan (1994) and Glosten, Jaganathan, and Runkle (1993). The set of models is comprised of

Symmetric models

| $\operatorname{ARCH}(1)-\mathrm{M}$ and | $\operatorname{ARCH}(2)-\mathrm{M}$ |  |
| :--- | :--- | :--- |
| $\operatorname{GARCH}(1,1)-\mathrm{M}$ | $\operatorname{GARCH}(2,1)-\mathrm{M}$ | $\operatorname{GARCH}(1,2)-\mathrm{M}$ and $\quad \operatorname{GARCH}(2,2)-\mathrm{M}$ |

Asymmetric models

| $\operatorname{TARCH}(1)-\mathrm{M}$ | $\operatorname{TARCH}(2)-\mathrm{M}$ |  |  |
| :--- | :--- | :--- | :--- |
| $\operatorname{TARCH}(1,1)-\mathrm{M}$ | $\operatorname{TARCH}(2,1)-\mathrm{M}$ | $\operatorname{TARCH}(1,2)-\mathrm{M}$ and | $\operatorname{TARCH}(2,2)-\mathrm{M}$ |
| $\operatorname{EARCH}(1)-\mathrm{M}$ | $\operatorname{EGARCH}(1,1)-\mathrm{M}$ and | $\operatorname{EGARCH}(2,1)-\mathrm{M}$ |  |

In table 2 the results of these conditional variance models chosen are reported.
To better understand this table, it is necessary to expose our dating convention.

| Mean equation |  |
| :---: | :---: |
|  | $m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{U}}=\phi\left(\sigma_{\mathrm{t}}^{\mathrm{U}}\right)^{2}+\varepsilon_{\mathrm{t}+1}^{\mathrm{i}}$ |
| Variance equation |  |
| Symmetric models: |  |
| ARCH (q) model | $\left(\sigma_{\mathrm{t}}^{\mathrm{U}}\right)^{2}=\varpi+\sum_{\mathrm{j}=1}^{\mathrm{q}} \delta_{\mathrm{q}}\left(\varepsilon_{\mathrm{t}+1-\mathrm{q}}^{\mathrm{U}}\right)^{2}$ |
| GARCH ( $\mathrm{p}, \mathrm{q}$ ) model | $\left(\sigma_{\mathrm{t}}^{\mathrm{U}}\right)^{2}=\varpi+\sum_{\mathrm{i}=1}^{\mathrm{p}} \eta_{\mathrm{i}}\left(\sigma_{\mathrm{t}-\mathrm{i}}^{\mathrm{U}}\right)^{2}+\sum_{\mathrm{j}=1}^{\mathrm{q}} \delta_{\mathrm{q}}\left(\varepsilon_{\mathrm{t}+1-\mathrm{q}}^{\mathrm{U}}\right)^{2}$ |
| Asymmetric models: | $\left(\sigma_{\mathrm{t}}^{\mathrm{U}}\right)^{2}=\varpi+\sum_{\mathrm{i}=1}^{\mathrm{p}} \eta_{\mathrm{i}}\left(\sigma_{\mathrm{t}-\mathrm{i}}^{\mathrm{U}}\right)^{2}+\gamma d_{\mathrm{t}}\left(\varepsilon_{\mathrm{t}}^{\mathrm{U}}\right)^{2}+\sum_{\mathrm{j}=1}^{\mathrm{q}} \delta_{\mathrm{q}}\left(\varepsilon_{\mathrm{t}+1-\mathrm{q}}^{\mathrm{U}}\right)^{2}$ |
| TARCH (p,q) |  |
| EGARCH (p,1) | $\ln \left[\left(\sigma_{\mathrm{t}}^{\mathrm{U}}\right)^{2}\right]=\varpi+\sum_{\mathrm{i}=1}^{\mathrm{p}} \theta_{\mathrm{i}} \ln \left[\left(\sigma_{\mathrm{t}-\mathrm{i}}^{\mathrm{U}}\right)^{2}\right]+\gamma_{1}$ |

where $\varphi_{\mathrm{t}} \equiv\left[\left(\sigma_{\mathrm{t}}^{\mathrm{C}}\right)^{2}-\left(\sigma_{\mathrm{t}}^{\mathrm{U}}\right)^{2}\right]$. According to (74), testing the the log-linearized asset pricing is equivalent to testing the null $\hat{\alpha}_{1}=0$ and $\hat{\psi}=0$.

We also run two other adjusted regressions which are simplified versions of (79). When we assume that $\hat{\psi}=0$, we are able to evidence the power of the forward premium to predict the term $s_{\mathrm{t}+1}-s_{\mathrm{t}}-0.5 \varphi_{\mathrm{t}}$, where the null hypothesis consists in testing $\hat{\alpha}_{1}=0$ in regression

$$
\begin{equation*}
s_{\mathrm{t}+1}-s_{\mathrm{t}}-0.5 \varphi_{\mathrm{t}}=\left(1-\alpha_{1}\right)\left(\mathrm{t}_{\mathrm{t}} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+\varepsilon_{\mathrm{t}+1} \tag{80}
\end{equation*}
$$

Finally, assuming $\hat{\alpha}_{1}=0$ we run the regression

$$
\begin{equation*}
s_{\mathrm{t}+1}-\mathrm{t} f_{\mathrm{t}+1}=(0.5-\psi) \varphi_{\mathrm{t}}+\varepsilon_{\mathrm{t}+1} \tag{81}
\end{equation*}
$$

which enables us to evidence the power of the omitted term to explain the term $s_{\mathrm{t}+1}-_{\mathrm{t}} f_{\mathrm{t}+1}$. In this case, the null hypothesis consists in testing $\hat{\psi}=0$. The results of all these estimates are reported in table 3 .

## 4. Empirical Results

### 4.1. Data and Summary Statistics

In principle, whenever econometric or statistical tests are performed, it is preferable to employ a long data set either in the time-series or in the cross-sectional dimension. There are some obvious limitations, especially when the FPP is concerned, the main one being due to the forward rate series, since a sufficiently large time-series is hardly available ${ }^{53}$. In order to have a common sample we covered the period starting in 1977:1 and ending in 2001:3, with quarterly frequency. We collected foreign exchange data for the following countries: Canada, Germany, Japan, Switzerland, U.K. and the U.S.

To analyze the FPP, we need first to compute the US\$ real returns on short-term British, Canadian, German, Swiss and Japanese government bonds, where both spot and forward exchange-rate data were used to transform returns denominated in foreign currency into US\$. The forward rate series were extracted from the Chicago Mercantile of Exchange database, while the spot rate series were extracted from Bank of England database.

A second ingredient for testing this puzzle is the construction of a return-based estimate of the pricing kernel ( $M_{\mathrm{t}}^{*}$ ). In this sense we follow da Costa et al. (2006) using their global data set, mostly composed of aggregate returns on stocks and government bonds for G7 countries. It covers US\$ real returns on G7-country stock indices and short-term government bonds, where spot exchange rate data were used to transform returns denominated in foreign currency into US $\$$ and the consumer price index of services and nondurable goods in the US was used as a deflator. In addition to G7 returns on stocks and bonds, it also contains US\$ real returns on gold, US real estate, bonds on AAA US corporations, Nasdaq, Dow Jones and the S\&P500. The US government bond is chosen to be the 90-day T-Bill. Regarding data sources, the returns on G7 government bonds were extracted from IFS/IMF, the returns on Nasdaq and Dow Jones Composite Index were extracted from Yahoo finance, the returns on real-estate trusts were extracted from the National Association of Real-Estate Investment Trusts in the US ${ }^{54}$, while the remaining series were extracted from the DRI database. Our sample period starts in 1977:1 and ends in 2001:3, with quarterly frequency. These portfolios of assets cover a wide spectrum of investment

[^30]opportunities across the globe, which is an important element of our choice of assets, since diversification is recommended for both methods of computing $M_{\mathrm{t}}^{*}$.

Table 1 presents summary statistics of our database over the period 1977:1 to 2001:3. Except for the Swiss case, the real return on covered trading of foreign bonds show means and standard deviations quite similar to that of the U.S. Treasury Bill, which is reasonable, since it reflects the covered interest rate parity in a frictionless economy. Regarding the return on uncovered trading, the means are ranging from $1.95 \%$ to $4.30 \%$, while the standard deviation from $4.90 \%$ to $16.84 \%$. Regarding the return on uncovered trading, means range from $2.47 \%$ to $4.70 \%$, while standard deviations range from $5.44 \%$ to $16.18 \%$.

### 4.2. SDF Estimate

The data set to estimate $M_{\mathrm{t}}^{*}$ makes heavily use of a reduced number of global portfolios. Despite this reduced number, it is plausible to expect the operation of a weak-law of large numbers, since these portfolios contain many assets. For example, we use aggregate returns on equity for G7 countries. For each country, these portfolios contain hundreds of asset (sometimes more than a thousand, depending on the country being considered). Therefore, the final average containing these portfolios will be formed using thousands of assets worldwide. Of course, it would be preferable to work at a more disaggregated level, e.g., at the firm level. However, there is no comprehensive database with a long enough time span containing worldwide returns at the firm level. There is a trade-off between $N$ and $T$ for available databases, which today limits empirical studies.

It is important to note that we carefully select the assets used to estimate the SDF and the exclusion of potential assets from our sample does not mean that the agents did not trade on these assets. This is something often forgotten in studies involving the financial markets but the point is that the "parameter" $M_{\mathrm{t}}^{*}$ we want to estimate is not altered by our restricting the information we use in estimating it! The relevant question is whether we are capable of identifying the realized SDF from the observation of a sample of assets. The problem, therefore, is not one of what space is spanned by each set of asset, as one might have been lead to think, but whether the observed assets are representative in the sense of satisfying the assumptions presented in the previous subsections.

Estimate $\widetilde{M}_{\mathrm{t}}$ is plotted in Figure 1, which also includes their summary statistics. It is possible to evidence that the multifactor model generates a SDF such that: i) is more volatile than the one obtained from canonical consumption based models and $i$ i) its mean is in accordance with the literature, since the respective unconditional risk free rate is $2,03 \%$ annualized.

It is worth mentioning the excellent performance of $\widetilde{M}_{\mathrm{t}}$ in the pricing tests proposed by da Costa et al. (2006). According to their results, it is not possible to reject the hypothesis that the pricing kernel thus constructed prices correctly all returns and excess returns on foreign assets as well as the excess return on equity over short term risk-free bonds in the US and the return on this risk-free bond, providing an evidence that this pricing kernel accounts for both the Forward and the Equity premium puzzles.

Finally, when the multi-factor model is used, since factors may be viewed as traded portfolios, it is interesting to examine which assets carry the largest weights in the SDF composition. For this global data set, in order to specify the three factors used, the assets with largest weights are the German, British and American stock indices, the German, Japanese and British government bonds and Nasdaq.

### 4.3 Omitted term

Initially, we present the estimate of the omitted term $\varphi_{\mathrm{t}}$.
According to our results reported in table 2, where we show the conditional variance models chosen for both process $m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{U}}$ and $m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{C}}$ considering each foreign currency, for the covered trading of Canadian and Japanese government bonds and for both covered and uncovered trading of British government bonds we reject the hypothesis of absence of asymmetric effects on the variance equation of the process, which limits our analyses to the asymmetric models in accordance with the Akaike and Schwarz criteria information. For the remaining process, it is not possible to reject the absence of any "sign" of asymmetric effects on the variance equation, limiting our choice to the symmetric models based on the same criteria information.

Regarding the mean equation, one can note that for every process we fail to reject the hypothesis that the conditional variance does not affect the mean, an uncomfortable result since it would not be expected in this log-linearized framework.

To frame our discussion about the relevance of the inclusion of the omitted term, table 3 presents estimates of the conventional and the adjusted regressions.

With regards the conventional regression, we evidence the disappointing forward bias for the German, Japanese and Swiss cases. For the Canadian and the British cases, although the values of the slope coefficient can be considered unreasonable, they are closely significant at $10 \%$ significance level, but not significant at $5 \%$ level.

Considering that these results work as our benchmark, when we run the adjusted regression (79), the results suggest that for the Japanese case the inclusion of this omitted term helps (but it is not enough) to accommodate the forward bias. For the remaining foreign currencies, this inclusion does not change the value nor the significance of the slope coefficients.

Regarding the adjusted regression (80), except for the Swiss case, the inclusion of the $-\frac{1}{2}\left[\left(\sigma_{\mathrm{t}}^{\mathrm{C}}\right)^{2}-\left(\sigma_{\mathrm{t}}^{\mathrm{U}}\right)^{2}\right]$ term in the left hand side of the conventional regression seems to be useful in reducing considerably the value of the coefficient $\alpha_{1}$, but once more not enough to provide the "correct" forward predictability power.

Finally, according to the results of the adjusted regression (81), even for the Japanese case where the coefficient $\psi$ assumes an undesirable value, the inclusion of the omitted term seems to account for the foreign currency risk premium.

To summarize. Although the omitted term "explains" the risk premium in accordance with the log-linearized asset pricing equation, it has a poor performance when of its inclusion in the conventional regression without imposing any restriction about the explanatory power of the forward premium.

One must observe that the "negative" side of our results suggests the rejection of the joint hypothesis that we are working with an appropriate risk model and also that the assumption 3 holds, justifying the implementation of the test of the lognormality of the process $m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{U}}$ and $m_{\mathrm{t}+1}+r_{\mathrm{t}+1}^{\mathrm{C}}$ for each currency used.

Besides this lognormality test, in the forthcoming extension of this working paper, we also intend to analyse the econometric issues of the conventional and the adjusted regressions proposed here, performing: i) formal tests of the stability of the coefficients ${ }^{55}$ and $\left.i i\right)$ tests of stationarity and cointegration of the appropriate variables.

[^31]Finally we still intend to analyse the individual role played by the $\operatorname{cov}_{\mathbf{t}}\left(s_{\mathbf{t}+1}, m_{\mathbf{t}+1}-\pi_{\mathbf{t}+1}\right)$ term in the equation (75) using a multivariate $\operatorname{GARCH}(\mathrm{p}, \mathrm{q})-\mathrm{M}$ model, following the approaches proposed by Goeij and Marquering (2002) and Bollerslev et al. (1992).

## 5. Conclusions

The FPP can be characterized by a quest for a theoretically sound economic model able to provide a definition of risk capable of correctly pricing the forward premium. But, even though one can write this model, the robust and disappointing evidence that domestic currency is expected to appreciate when domestic nominal interest rates exceed foreign interest rates remains to be accommodated.

Aiming to better understand this uncomfortable situation, we work directly and only with the log-linearized Asset Pricing Equation, which enables us to characterize not only the necessary conditions but also the sufficient ones that the SDF must display if it is to account for the FPP.

We then propose and test a novel three-stage approach, running adjusted regressions that have the conventional one as a particular case. Our results suggest that although the omitted term "explains" the risk premium in accordance with the log-linearized asset pricing equation, its inclusion in the conventional regression used to evidence the puzzle does not change the significance nor the magnitude of the strong predictability power of the forward premium.

## References

[1] Araujo, F., Issler, J.V. and Fernandes, M. (2005), "Estimating the stochastic discount factor without a utility function," Working Paper \# 583, Graduate School of Economics, Getulio Vargas Foundation, downloadable as http://epge.fgv.br/portal/arquivo/1817.pdf
[2] Backus, D., Foresi, S. and Telmer, C. (1995), "Interpreting the forward premium anomaly," The Canadian Journal of Economics 28, 108-119.
[3] Bai, J. (2005), "Panel Data Models with Interactive Fixed Effects," Working Paper: New York University.
[4] Bailie, R. T. and Bollerslev, T. (2000), "The forward premium anomaly is not as bad as you think," Journal of International Money and Finance 19, 471-488.
[5] Bekaert, G. (1996), "The time variation of risk and return in foreign exchange markets: a general equilibrium perspective," Review of Financial Studies 2, 427-470.
[6] Bekaert G. and Hodrick, R. J. (1993), "On biases in the measurement of foreign exchange risk premiums," Journal of International Money and Finance 12, 115-138.
[7] Bollerslev, T., Chou, R. Y. and Kroner, K. F. (1992), "ARCH modelling in finance: a review of theory and empirical evidence", Journal of Econometrics 52, 5-59.
sample sizes and very persistent autocorrelation in the forward premium.
[25] Froot, K. A. (1990), "Short rates and expected asset returns," NBER Working paper n ${ }^{0} 3247$.
[26] Goeij, P. D. and Marquering, W. (2002), "Modeling the conditional covariance between stock and bond returns: a multivariate GARCH approach", Mimeo. Erasmus University Roterdam.*
[27] Groen, J. J. and Balakrishnan, R. (2006),"Asset price based estimates of stern014359(bas)-1(ed)-b09.6Age rate(ba premia", Journal of I9.nternational Money and Finane5, 71-92.
[28] Hansen, L. P. and Hodrick, R. J.
[41] Mehra, R. and Prescott, E. (1985), "The equity premium: a puzzle," Journal of Monetary Economics 15, 627-636.
[42] Mulligan, C. (2002), "Capital, interest, and aggregate intertemporal substitution," Working Paper \# w9373: National Bureau of Economic Research.
[43] Nelson, D. (1991), "Conditional Heteroskedasticity in asset returns: a new approach", Econometrica 59 (2), 347-370.
[44] Ross, S.A. (1976), "The arbitrage theory of capital asset pricing", Journal of Economic Theory, 13, pp. 341-360.
[45] West, K. D. and Cho, D. (1995), "The predictability of several models of exchange rate volatility", Journal of Econometrics, 69, pp. 367-391.
Tables and Figures
Table 1: Data summary statistics

| International quarterly data: observations from 1977:I to 2001:III |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\begin{array}{r} U S \$ \text { real } \mathrm{r} \\ \text { trading of for } \end{array}$ | turn on the covered <br> ign government bonds | $\begin{gathered} U S \$ \text { real re } \\ \text { trading of for } \end{gathered}$ | urn on the uncovered ign government bonds | $\begin{gathered} U S \$ \text { real } \\ \text { uncovered o } \\ \text { of foreigr } \end{gathered}$ | xcess return on the er the covered trading government bonds |
| Government <br> Bonds | Sample mean <br> (\% per year) | Sample stand. deviat. <br> (\% per year) | Sample mean (\% per year) | Sample stand. deviat. <br> (\% per year) | Sample mean (\% per year) | Sample stand. deviat. <br> (\% per year) |
| Canadian | 2.612 | 2.232 | 1.950 | 4.902 | -0.661 | 4.187 |
| Deutsch | 2.941 | 1.914 | 2.980 | 11.706 | -0.961 | 11.806 |
| Japanese | 3.021 | 1.729 | 4.299 | 13.351 | 1.279 | 13.283 |
| Swiss | 8.712 | 16.220 | 2.682 | 16.842 | -6.029 | 20.300 |
| British | 2.030 | 2.263 | 4.264 | 10.330 | 2.233 | 10.631 |


Table 2: Models of conditional variance

Table 2 (continued)

| Swiss government bonds |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\hat{\phi}$ | $\hat{\varpi}$ | $\hat{\eta}_{1}$ | $\hat{\delta}_{1}$ |  |
| $\sigma_{\mathrm{t}}^{2}\left[\ln \left(\widetilde{M_{\mathrm{t}+1}} R_{\mathrm{t}+1}^{\mathrm{C}}\right)\right]$ | $\operatorname{GARCH}(1,1)-\mathrm{M}^{1 ; 2}$ | 0.053 | 0.020 | 0.646 | -0.111 |  |
|  |  | (0.915)* | $(0.190)^{*}$ | (0.052)* |  |  |
|  |  | $\hat{\phi}$ | $\hat{\varpi}$ | $\hat{\eta}_{1}$ | $\hat{\delta}_{1}$ | $\hat{\delta}_{2}$ |
| $\sigma_{\mathrm{t}}^{2}\left[\ln \left(M_{\mathrm{t}+1} R_{\mathrm{t}+1}^{\mathrm{U}}\right)\right]$ | $\operatorname{GARCH}(1,2)-\mathrm{M}^{2}$ | -0.283 | -0.003 | 1.062 | -0.112 | 0.099 |
|  |  | $(0.542)^{*}$ | (0.000) | (0.000) | (0.001) | (0.005) |

British government bonds

| $\sigma_{\mathrm{t}}^{2}\left[\ln \left(\widetilde{M_{\mathrm{t}+1}} R_{\mathrm{t}+1}^{\mathrm{C}}\right)\right]$$\left.\sigma_{\mathrm{t}}^{2}\left[\ln \widetilde{\left(M_{\mathrm{t}+1}\right.} R_{\mathrm{t}+1}^{\mathrm{U}}\right)\right]$ | $\operatorname{TARCH}(1,1)-\mathrm{M}^{1 ; 2}$ | $\hat{\phi}$ |  |  | $\hat{\eta}_{1}$ | $\hat{\gamma}$ | $\hat{\delta}_{1}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | -0.427 |  |  | 0.533 | 0.202 | 0.057 |
|  |  | $(0.381)^{*}$ |  |  | $(0.168)^{*}$ | (0.000) | (0.000) |
|  | TARCH $(2,1)-\mathrm{M}^{1}$ | $\hat{\phi}$ | $\hat{\varpi}$ | $\hat{\eta}_{1}$ | $\hat{\eta}_{2}$ | $\hat{\gamma}$ | $\hat{\delta}_{1}$ |
|  |  | -0.347 | 0.020 | 1.311 | -0.676 | -0.113 | 0.014 |
|  |  | $(0.459)^{*}$ | (0.036) | (0.000) | (0.015) | $(0.061)^{*}$ | $(0.716)^{*}$ |

Table 3: Conventional and adjusted regressions
$\begin{aligned} \text { Notes: } & { }^{1} \text { OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987). } \\ & { }^{*} \text { Reject the null hypothesis at the } 5 \% \text { significance level. P- values in the parenthesis. }\end{aligned}$
Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).
${ }^{*}$ Reject the null hypothesis at the $5 \%$ significance level. P- values in the parenthesis.

| Conventional regression: | Canadian |  | Foreign government bonds |  |  |  |  |  | British |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | German |  | Japanese |  | Swiss |  |  |  |
|  | $\hat{\alpha}_{0}$ | $\hat{\alpha}_{1}$ | $\hat{\alpha}_{0}$ | $\hat{\alpha}_{1}$ | $\hat{\alpha}_{0}$ | $\hat{\alpha}_{1}$ | $\hat{\alpha}_{0}$ | $\hat{\alpha}_{1}$ | $\hat{\alpha}_{0}$ | $\hat{\alpha}_{1}$ |
| $s_{\mathrm{t}+1}-s_{\mathrm{t}}=\alpha_{0}+$ | 0.003 | 0.472 | 0.004 | 1.713 | 0.028 | 3.257 | -0.002 | 0.748 | -0.006 | 1.649 |
| $+\left(1-\alpha_{1}\right)\left(\mathrm{t} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+\varepsilon_{\mathrm{t}+1}$ | $(0.173)^{*}$ | $(0.072)^{*}$ | $(0.520)^{*}$ | (0.007) | (0.003) | (0.000) | $(0.814)^{*}$ | (0.000) | $(0.397) *$ | $(0.118){ }^{*}$ |
| Adjusted regressions: |  |  |  |  |  |  |  |  |  |  |
|  | $\hat{\psi}$ | $\hat{\alpha}_{1}$ | $\hat{\psi}$ | $\hat{\alpha}_{1}$ | $\hat{\psi}$ | $\hat{\alpha}_{1}$ | $\hat{\psi}$ | $\hat{\alpha}_{1}$ | $\hat{\psi}$ | $\hat{\alpha}_{1}$ |
| $s_{\mathrm{t}+1}-s_{\mathrm{t}}=(0.5-\psi) \varphi_{\mathrm{t}}+$ | 0.113 | 0.282 | 0.712 | 1.635 | 0.574 | 1.467 | 0.471 | 0.753 | 1.583 | 1.824 |
| $+\left(1-\alpha_{1}\right)\left({ }_{\mathrm{t}} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+\varepsilon_{\mathrm{t}+1}$ | $(0.587)^{*}$ | $(0.255)^{*}$ | $(0.143)^{*}$ | (0.009) | $(0.106)^{*}$ | (0.017) | $(0.060)^{*}$ | (0.000) | (0.002) | (0.017) |
|  | $\hat{\alpha}_{1}$ |  | $\hat{\alpha}_{1}$ |  | $\hat{\alpha}_{1}$ |  | $\hat{\alpha}_{1}$ |  | $\hat{\alpha}_{1}$ |  |
| $s_{\mathrm{t}+1}-s_{\mathrm{t}}-0.5 \varphi_{\mathrm{t}}=$ | 0.261 |  | 1.205 |  | 1.189 |  | 0.711 |  | 1.233 |  |
| $=\left(1-\alpha_{1}\right)\left(\mathrm{t} f_{\mathrm{t}+1}-s_{\mathrm{t}}\right)+\varepsilon_{\mathrm{t}+1}$ | $(0.284) *$ |  | (0.029) |  | (0.044) |  | (0.001) |  | $(0.183){ }^{*}$ |  |
|  | $\hat{\psi}$ |  | $\hat{\psi}$ |  | $\hat{\psi}$ |  | $\hat{\psi}$ |  | $\hat{\psi}$ |  |
| $s_{\mathrm{t}+1}-\mathrm{t} f_{\mathrm{t}+1}=$ | 0.076 |  | -0.104 |  | 0.332 |  | 0.039 |  | 0.068 |  |
| $=(0.5-\psi) \varphi_{\mathrm{t}}+\varepsilon_{\mathrm{t}+1}$ | $(0.712)^{*}$ |  | $(0.828){ }^{*}$ |  | $(0.340)^{*}$ |  | $(0.911)^{*}$ |  | $(0.900){ }^{*}$ |  |

# Livros Grátis 

( http://www.livrosgratis.com.br )
Milhares de Livros para Download:
Baixar livros de Administração
Baixar livros de Agronomia
Baixar livros de Arquitetura
Baixar livros de Artes
Baixar livros de Astronomia
Baixar livros de Biologia Geral
Baixar livros de Ciência da Computação
Baixar livros de Ciência da Informação
Baixar livros de Ciência Política
Baixar livros de Ciências da Saúde
Baixar livros de Comunicação
Baixar livros do Conselho Nacional de Educação - CNE
Baixar livros de Defesa civil
Baixar livros de Direito
Baixar livros de Direitos humanos
Baixar livros de Economia
Baixar livros de Economia Doméstica
Baixar livros de Educação
Baixar livros de Educação - Trânsito
Baixar livros de Educação Física
Baixar livros de Engenharia Aeroespacial
Baixar livros de Farmácia
Baixar livros de Filosofia
Baixar livros de Física
Baixar livros de Geociências
Baixar livros de Geografia
Baixar livros de História
Baixar livros de Línguas
Baixar livros de Literatura
Baixar livros de Literatura de Cordel
Baixar livros de Literatura Infantil
Baixar livros de Matemática
Baixar livros de Medicina
Baixar livros de Medicina Veterinária
Baixar livros de Meio Ambiente
Baixar livros de Meteorologia
Baixar Monografias e TCC
Baixar livros Multidisciplinar
Baixar livros de Música
Baixar livros de Psicologia
Baixar livros de Química
Baixar livros de Saúde Coletiva
Baixar livros de Serviço Social
Baixar livros de Sociologia
Baixar livros de Teologia
Baixar livros de Trabalho
Baixar livros de Turismo


[^0]:    ${ }^{1}$ First draft July 2005. We thank Caio Almeida, Marco Bonomo, Luis Braido, Marcelo Fernandes, Jaime Filho, Luis Renato Lima, Celina Ozawa and participants of the XXVII Meeting of the Brazilian Society of Econometrics, the "Econometrics in Rio" meeting, the 2006 Meeting of the Brazilian Finance Society and the $61^{\text {th }}$ European Meeting of the Econometric Society for their comments and suggestions. The usual disclaimer applies.
    ${ }^{2}$ We do not claim that returns on equity are not predictable. In fact we know that book-to-market ratios and other variables "predict" returns. The point is that this empirical regularity was never seen as defining the EPP.

[^1]:    ${ }^{3}$ In what follows, we use the terms pricing kernel and stochastic discount factor interchangeably.

[^2]:    ${ }^{4}$ See the comprehensive surveys by Hodrick (1987) and Engel (1996).
    ${ }^{5}$ In what follows we use capital letters to denote variables in levels and small letters to denote the logs of these variables.
    ${ }^{6}$ As we shall see later, weaker, though still restrictive, conditions on the behavior of risk premium may be compatible with $\alpha_{1}=1$.

[^3]:    ${ }^{7}$ Frankel (1979), for example, argues that most exchange rate risks are diversifiable, there being no grounds for agents to be rewarded for holding foreign assets.
    ${ }^{8}$ The findings in the case of EPP are similar, in the sense that in both cases the estimated values are above 10, which is viewed as unacceptable and unreasonable, according to Krugman (1981), Mehra and Prescott (1985) and Romer (1995).

[^4]:    ${ }^{11}$ We should also emphasize that we do not reject (9) for any of the instruments, as well, which means that our SDF satisfies both conditions presented by Backus et al. (1995).

[^5]:    ${ }^{12}$ Since, according to the literature, $\left(s_{t+1}-s_{t}\right)^{\sim} I(0)$ and $s_{t} \sim I(1)$, one must be careful when defining the $\mathbb{V}_{t}\left(s_{t+1}\right)$ and $\operatorname{cov}_{t}\left(s_{t+1}, m_{t+1}-\pi_{t+1}\right)$ terms. Otherwise one may end up with an unbalanced regression.
    ${ }^{13}$ For more details about the heteroskedasticity of asset returns, see Bollerslev et al. (1988), Engle et al. (1990) and Engle and Marcucci (2005).

[^6]:    ${ }^{15}$ Chicago Mercantile Exchange pioneered the development of financial futures with the launch of currency futures, the world's first financial futures contracts, in 1972.
    ${ }^{16}$ Data on the return on real estate are measured using the return of all publicly traded REITs - Real-Estate Investment Trusts.

[^7]:    ${ }^{17}$ In order to obtain (16) one needs a large negative covariance between $M_{t}$ and $\left(i_{t}^{S P}-i_{t-1}^{b}\right) P_{t-1} / P_{t}$. This happens because $\mathbb{E}\left(M_{t}\right) \cong 1$, therefore, (16) implies:

    $$
    -\operatorname{cov}\left(M_{t}, \frac{\left(i_{t}^{S P}-i_{t-1}^{b}\right) P_{t-1}}{P_{t}}\right) \cong \mathbb{E}\left\{\frac{\left(i_{t}^{S P}-i_{t-1}^{b}\right) P_{t-1}}{P_{t}}\right\}
    $$

    Notice that $\mathbb{E}\left\{\left(i_{t}^{S P}-i_{t-1}^{b}\right) P_{t-1} / P_{t}\right\}$ is the mean equity premium, a large number by all accounts. Therefore, (16) requires first a negative covariance between $M_{t}$ and $\left(i_{t}^{S P}-i_{t-1}^{b}\right) P_{t-1} / P_{t}$, which must be sufficiently large in absolute value.

[^8]:    ${ }^{18}$ This position should however be contrasted to that implicit in Brandt et al (2006), where the behavior of the SDF is viewed as being equal to that of the marginal rate of substitution for a model of preferences and/or market structure yet to be written.
    ${ }^{19}$ A good example of evidence against the simple representative consumer framework is the comparision between the time-series literature on optimal consumption and the panel-data literature. While in the first, consumption-based estimates using aggregate consumption generated important puzzles in finance and in macroeconomics, such as excess smoothness (excess sensitivity) in consumption, the equity-premium puzzle and the forward-premium puzzle - Hansen and Singleton (1982, 1983, 1984), Mehra and Prescott (1985), Mark (1985), Campbell (1987), Campbell and Deaton (1989), Hodrick (1989), Epstein and Zin (1991), and Engel (1996) - the second shows very little evidence of these puzzles - Runkle (1991), Blundell, Browning, and Meghir (1994), Attanasio and Browning (1995), and Attanasio and Weber (1995). One possible explanation for these findings is inadequate crosssectional aggregation of the euler equation, which is a non-linear relationship. Recently, the work of Mulligan (2002, 2004) offers a reconciliation of time-series and panel-data results.

[^9]:    standard deviations are multiplied by 200 .

[^10]:    Notes: ${ }^{1}$ OLS estimate using the heteroskedasticity and autocorrelation consistent covariance estimator proposed by Newey and West (1987).

    * Reject the null hypothesis at the $5 \%$ significance level.

[^11]:    ${ }^{20}$ First draft November 2006. We thank Caio Almeida, Luis Braido, Luis Renato Lima and seminar participants at Getulio Vargas Foundation for their comments and suggestions. The usual disclaimer applies.
    ${ }^{21}$ See the comprehensive surveys by Hodrick (1987) and Engel (1996).

[^12]:    ${ }^{24}$ See e.g. Samuelson (1969), Rubinstein (1976), Lucas (1978), Breeden (1979), Grossman and Shiller (1981) and some other classical macroeconomic and financial papers.
    ${ }^{25}$ Here, we are implicitly assuming the absence of short-sale cosntraints besides other frictions in the economy, despite we recognize the significance of bid-ask spreads' impact on the profitability of currency speculation, as mentioned in Burnside et al. (2006).

[^13]:    ${ }^{26}$ Much of the volatility of the return on 90 -day Treasury bill is certainly due to short-run inflation risk.

[^14]:    ${ }^{27}$ Cochrane (2001) argues that depending on the state variable and on the utility function, the choice of an external or internal habit model can be seen as a technical convenience, affecting slightly the results.
    ${ }^{28}$ There might not be a clear answer to the question of which one is optimal for each case, since both models can be justified empirically.

[^15]:    ${ }^{29}$ Alvarez and Jermann's (2002) argues that using SDF as a function of only consumption data it is not possible to provide a permanent component term higher than the lower bound.

[^16]:    ${ }^{30}$ The Campbell and Cochrane (1999) model can also accommodate more complex consumption processes, including processes with predictability, conditional heteroskedasticity, and nonnormality.

[^17]:    ${ }^{31}$ It is convenient, but not necessarily, to use the same value $g$ for the mean consumption growth rate and the parameter $g$ in the habit accumulation equation (53).
    ${ }^{32}$ This means that a higher GDP does not make agents happier if they do not have their incomes increase more than average.

[^18]:    ${ }^{33}$ Chicago Mercantile Exchange pioneered the development of financial futures with the launch of currency futures, the world's first financial futures contracts, in 1972.

[^19]:    ${ }^{35}$ One could argue that this empirical failure can associated with the problem of a near nonidentification of the risk aversion parameter in the canonical model.

[^20]:    ${ }^{36}$ The strategy adopted in these papers is different of ours, since the Euler equations derived from the Epstein and Zin (1989) specification for the excess return on equity over 90 -day T-bill and on the return on 90 -day T-bill are estimated jointly.

[^21]:    ${ }^{37}$ Not necessarily these non-rejections are much informative, since one could attribute them to the poor properties of conventional

[^22]:    ${ }^{38}$ Frankel (1979), for example, argues that most exchange rate risks are diversifiable, there being no grounds for agents to be rewarded for holding foreign assets.

[^23]:    ${ }^{42}$ See the comprehensive surveys by Hodrick (1987) and Engel (1996).
    ${ }^{43}$ In what follows we use capital letters to denote variables in levels and small letters to denote the logs of these variables.
    ${ }^{44}$ As we shall see later, weaker, though still restrictive, conditions on the behavior of risk premium may be compatible with $\alpha_{1}=0$.

[^24]:    ${ }^{45}$ Here, we are implicitly assuming the absence of short-sale cosntraints besides other frictions in the economy, despite we recognize the significance of bid-ask spreads' impact on the profitability of currency speculation, as mentioned in Burnside et al. (2006).

[^25]:    ${ }^{46}$ Here, we are implicitly assuming that this is a frictionless economy.

[^26]:    ${ }^{47}$ Here, $\mathbb{E}_{t}^{*}(\cdot)$ denotes the conditional expectation given the forward premium, ${ }_{t} f_{t+1}-s_{t}$.
    ${ }^{48}$ ince, according to the literature, $\left(s_{t+1}-s_{t}\right) \sim I(0)$ and $s_{t} \sim I(1)$, one must be careful when defining the $\mathbb{V}_{t}\left(s_{t+1}\right)$ and $\operatorname{cov}_{t}\left(s_{t+1}, m_{t+1}-\pi_{t+1}\right)$ terms. Otherwise one may end up with an unbalanced regression.

[^27]:    ${ }^{49}$ For more details about the heteroskedasticity of asset returns, see Bollerslev et al. (1988), Engle et al. (1990) and Engle and Marcucci (2005).

[^28]:    ${ }^{50}$ One should note that we are not assuming the existence of a risk free rate. If this is the case, rather then estimate the intercept

[^29]:    ${ }^{51}$ See Bollerslev et al. (1992) for an excellent survey about conditional variance models.
    ${ }^{52}$ See, for exmaple, West and Cho (1995) and Malmsten and Terasvirta (2004).

[^30]:    ${ }^{53}$ Chicago Mercantile Exchange pioneered the development of financial futures with the launch of currency futures, the world's first financial futures contracts, in 1972.
    ${ }^{54}$ Data on the return on real estate are measured using the return of all publicly traded REITs - Real-Estate Investment Trusts.

[^31]:    ${ }^{55}$ However, also according to Bekaert and Hodrick (1993) the (1) have changed over time and observing more recent studies regarding this regression, it is possible to identify "problems" related to the empirical evidences of this predictability. For instance, according to Bailie and Bollerslev (2000), this anomaly may be viewed mainly as a statistical phenomenum from having small

