Escola de Pós-Graduação em Economia -- EPGE

Fundação Getulio Vargas

Essays on the relationship between the Equity and the Forward Premium Puzzles

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

Aluno: Paulo Rogério Faustino Matos

Professor Orientador: Carlos Eugênio Ellery Lustosa da Costa

Rio de Janeiro

2006

Escola de Pós-Graduação em Economia -- EPGE

Fundação Getulio Vargas

Essays on the Relationship between the Equity and the Forward Premium Puzzles

Tese submetida à Escola 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)

Caio Ibsen de Almeida (EPGE)

Marco Antônio Bonomo (EPGE)

Marcelo Fernandes (Queen Mary University of London)

Ricardo Dias de Oliveira Brito (Ibmec-SP)

Rio de Janeiro

2006

Contents

* Uma versão anterior deste capítulo circulou em Encontros acadêmicos sob o título “Do Equity and the Foreign Currency Risk Premiums Display Common Patterns?”

Acknowledgements

Agradeço,
A Deus pela 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 em que 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 tese e
Ao corpo docente, discente e demais funcionários da EPGE, instituição a qual serei eternamente grato.

List of Tables

Chapter 1

Chapter 2

Chapter 3

List of Figures

Chapter 1

Chapter 2

Chapter 3

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— EPP— and 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 Fabrício Linhares) 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 relationship between the equity and the forward premium puzzles. We construct return-based stochastic discount factors using only American domestic assets and check whether they price correctly the foreign currency risk premia. We avoid log-linearizations by using moments restrictions associated with euler equations. Our pricing kernel accounts for both domestic and international markets stylized facts that escape consumption based models. In particular, we fail to reject the null hypothesis that the foreign currency risk premium has zero price when the instrument is the own current value of the forward premium.

JEL Code: G12; G15. Keywords: Equity Premium Puzzle, Forward Premium Puzzle, Return-based Pricing Kernel.

1. Introduction

The Forward Premium Puzzle – henceforth, FPP – is how one calls the systematic departure from the intuitive proposition that the expected return to speculation in the forward foreign exchange market should be zero, conditional on available information.

One of the most acknowledged puzzles in international finance, the FPP was, in its infancy, investigated by Mark (1985) within the framework of the consumption capital asset pricing model – CCAPM. Perhaps, following Hansen and Singleton's (1982, 1984) earlier method, Mark used a non-linear GMM approach, which revealed the model's inability, in its canonical version, to account for its implicit over-identifying restrictions. The results found by Mark are similar to the ones found by Hansen and Singleton with respect to the equity premium, an idea carried forward by Mehra and Prescott (1985) who went on to propose what they have labelled the Equity Premium Puzzle – henceforth, EPP.

At first sight, it may seem surprising that such similar results were never properly linked, and we may only conjecture why the literature on the FPP and the EPP drifted apart after the work of Mark. We list two alternative explanations. First, the failure of the CCAPM was a great disappointment for the profession, since it meant the absence of a fully specified economic model that could price assets. Such unheartening finding may have lead to a momentary halt in research linking the equity- and the forward-premium puzzles[2]. Second, 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[3] – may have led many to believe that even if the CCAPM was capable of accounting for the equity premium it would not solve the FPP; see Engel (1996).

Thinking deeper about these two puzzles, we are forced to conclude that proving that they are related is currently an impossible task, since it requires the existence of a consumption model generating a pricing kernel that properly prices assets, showing the shortcomings of previous models. Because we do not have such a proper model today, we cannot relate the EPP and the FPP within a CCAPM framework. This may explain why relating these two puzzles was not tried before and why two distinct research agendas involving them appeared over time.

As is well known, 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 has focused on adding state variables to standard consumption-based pricing kernels to change its behavior; see Epstein and Zin (1989), Constantinides and Duffie (1996) and Campbell and Cochrane (1999).

In this paper we revisit the FPP and the EPP and ask whether they deserve two distinct agendas, or whether they are but two symptoms of the same illness: the incapacity of existing consumption-based models to generate the implied behavior of *a* pricing kernel that correctly prices asset returns.

Given the limitations on proving that the FPP and the EPP are related, we use an indirect approach. If these two puzzles are solely a symptom of the inappropriateness of existing consumption-based pricing kernels, then they will not be manifest when appropriate pricing kernels are used. Suppose we find a single pricing kernel that is not a function of consumption and is compatible with all regularities in domestic financial markets not accounted for by current consumption-based kernels. At the same time, suppose that this pricing kernel accounts for the behavior of the forward premium. Then, we have good reason to believe that we should not disperse our research effort on two completely different agendas, but rather concentrate on a single one focused on rethinking consumption-based pricing kernels[4] – the common suspect for the two puzzles.

A crucial issue of our approach is to find what an *appropriate* pricing kernel is in this context. Hansen and Jagganathan (1991) have lead the profession towards return-based kernels instead of consumption-based kernels; see also Connor and Korajczyk (1986), Chamberlain and Rothschild (1983), and Bai (2005). The idea is to combine statistical methods with economic theory – the Asset Pricing Equation – to devise pricing-kernel estimates as the unique projection of a stochastic discount factor – henceforth, SDF – on the space of returns: the SDF mimicking portfolio. The latter can be estimated without any assumptions on a functional form for preferences, despite having a strong footing on theory as a consequence of the use of the Asset Pricing Equation.

One way to rationalize the SDF mimicking portfolio is to realize that it is the projection of a proper consumption model (yet to be written) on the space of payoffs. Thus, the pricing properties of this projection are no worse than those of the proper model – a key insight of Hansen and Jagganathan. An advantage of concentrating on the projection is that we can approximate it arbitrarily well in-sample using statistical methods and asset returns alone. Therefore, using such projection not only circumvents the inexistence of a proper consumption model but is also guaranteed not to underperform such ideal model.

Bearing in mind our stated goal, we extract a time series for the pricing kernel that does not depend on preferences or on consumption data. Two techniques are considered in this paper to estimate the SDF mimicking portfolio: i) Hansen and Jagganathan's mimicking portfolio, which is the projection of any stochastic discount factor on the space of returns and ii) the unconditional linear multifactor model, which is perhaps the dominant model in discrete-time empirical work in Finance.

As noted by Cochrane (2001), SDF estimates are just functions of data. Pricing correctly a specific group of assets can be achieved by building non-parsimonious SDF estimates, i.e., SDF estimates that price arbitrarily well that group of assets in-sample but not necessarily assets outside that group.[5] In order to avoid this critique, we construct SDF mimicking portfolio estimates using domestic (U.S.) returns alone – on the 200 most traded stocks in the NYSE, extracted from the CRSP database.

Assume we had a consumption-based model that did account for the equity premium.

Then, the behavior of its projection on the space of domestic-asset returns would have to coincide with that of an SDF mimicking portfolio built from these same assets. We then use this domestic SDF projection to price the forward premium. Tests are implemented for four countries within the G7 group, besides Switzerland, for which there exists a relatively long-span data for spot and future foreign exchange markets. Here, our tests show their out-of-sample character, avoiding Cochrane's critique.

Our tests make intensive use of the Asset Pricing Equation. They are all based on euler equations, exploiting theoretical lack of correlation between discounted risk premia and variables in the conditioning set, or between discounted returns and their respective theoretical means, i.e., we employ *discounted scaled excess-returns* and *discounted scaled returns* in testing. We investigate whether discounted risk premia have mean zero or whether discounted returns have a mean of unity.

Our results are clear cut: return-based pricing kernels using U.S. assets alone account for domestic stylized facts, pricing correctly the equity premium for the U.S. economy – which shows no signs of the EPP – and also pricing most of the Fama-French benchmark factor returns. At the same time, these same pricing-kernel estimates show no signs of the FPP in pricing the expected return to speculation in forward foreign-exchange markets for the widest group possible of developed countries with a long enough span of future exchange-rate data (Canada, Germany, Japan, Switzerland and the U.K.). This evidence raises the question of whether the FPP and the EPP are two symptoms of the same illness.

We summarize our empirical results as follows. First, the null of zero discounted excess returns on equities is not rejected even when potentially interesting forecasting variables are used as instruments. Second, for most countries, the moment restrictions associated with the euler equations (Asset Pricing Equations) are not rejected for excess returns and returns on operations with foreign assets for any of the instruments used. This includes the own current value of the forward premium, which shows no signs of predictability of the expected return to speculation, contradicting one of the defining features of the FPP. Only in the case of British bonds the results are, in some sense, conflicting. In some occasions, we reject the null hypothesis that the foreign currency risk premium has zero price.

Our results can be viewed as new evidence supporting the usefulness of reuniting the research agendas on the EPP and the FPP. Although we cannot claim that a consumption model that did account for the behavior of the equity premium would also price correctly the forward premium, we can claim that its projection on the space of domestic returns would.[6] In our view, this is as far as one can go today in showing that these two puzzles are related.

As argued above, we search not for a consumption model of the SDF, but simply for a procedure that identifies the SDF mimicking portfolio circumventing the fact that we still lack a good model for pricing risk or risk premia. Employing SDF mimicking portfolio estimates allows to test directly the pricing of risk or risk premia by using the theoretical restrictions associated with the Asset Pricing Equation. In our context, there is neither the need to specify a full model for preferences (consumption SDF) nor the need to perform a log-linearization of the Asset Pricing Equation in pricing tests. In that sense, we are able to isolate possible causes for rejection of theory (EPP and FPP) not isolated by the previous literature.

The remainder of the paper is organized as follows. Section 2 gives an account of the literature that tries to explain the FPP and is related to our current effort. Section 3 discusses the techniques used to estimate the SDF and the pricing tests are implemented in this paper. Section 4 presents the empirical results obtained in this paper. Concluding remarks are offered in Section 5.

2. A Critical Literature Review

Most studies[7] report the existence of the FPP through the finding that $\widehat{\alpha}_1$ is significantly smaller than zero when running the regression,

$$s_{t+1} - s_t = \alpha_0 + \alpha_1({}_tf_{t+1} - s_t) + u_{t+1}, \tag{1}$$

where s_t is the log of the exchange rate at time t, ${}_tf_{t+1}$ is the log of time t forward exchange rate contract and u_{t+1} is the regression error.[8] Notwithstanding the possible effect of Jensen inequality terms, testing the uncovered interest rate parity (UIP) is equivalent to testing the null $\alpha_1 = 1$ and $\alpha_0 = 0$, along with the uncorrelatedness of residuals from the estimated regression.

Although the null is rejected in almost all studies, it should be noted that α_1 not being equal to one ought not to be viewed as evidence of market failure or some form of irrationality and, *per se*, does not imply the existence of a 'puzzle' since the uncovered parity needs only to hold exactly in a world of risk-neutral agents, or if the return on currency speculation is not risky. The probable reason why these findings came to be called a puzzle was the magnitude of the discrepancy from the null: 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. This implies an expected domestic currency appreciation when domestic nominal interest rates exceed foreign interest rates, contrary to what is needed for the UIP to hold.

Log-linear regressions such as (1) have a long tradition in economics. As is well known, getting to (1) from first principles requires stringent assumptions, something that is usually overlooked when hypothesis testing is later performed using it. Next, we shall make explicit how strong the assumptions that underlie the null tested in (1) are. Our departing point is the Pricing Equation,

$$1 = \mathbb{E}_t \left[M_{t+1} R_{i,t+1} \right] \quad \forall i = 1, 2, \cdots, N. \tag{2}$$

where $R_{i,t+1}$ is the return of asset i and M_{t+1} is the pricing kernel or stochastic discount factor, SDF (e.g. Hansen and Jagganathan (1991)), a random variable that discount payoffs in such a way that their price is simply the discounted expected value.

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 sufficient to guarantee, through Riesz representation theorem, the existence of a SDF, M_{t+1}. Log linearizing (2) makes it is possible to justify regression (1), but not without unduly strong assumptions on the behavior of discounted returns.

Gomes and Issler (2007) criticize the empirical use of the log-linear approximation of the Pricing Equation (2) leading to (1). First, is the usual criticism that any hypothesis test using results of a log-linear regression is a joint test which includes the validity of the log-linearization being performed, i.e., includes an auxiliary hypothesis in testing. Therefore, rejection can happen if the null is true but the log-linearization is inappropriate. Second, they show that it is very hard to find appropriate instruments in estimating log-linear regressions such as (1), since, by construction, lagged variables are correlated with the error term.

To understand this latter point, consider a second-order taylor expansion of the exponential function around x, with increment h,

$$e^{x+h} = e^x + he^x + \frac{h^2 e^{x+\lambda(h)h}}{2}, \text{ with } \lambda(h) : \mathbb{R} \rightarrow (0,1)\,. \tag{3}$$

For a generic function, $\lambda(\cdot)$ depends on both x and h, but not for the exponential function. Indeed, dividing (3) by e^x, we get

$$e^h = 1 + h + \frac{h^2 e^{\lambda(h)h}}{2}, \tag{4}$$

showing that $\lambda(\cdot)$ depends only on h.[9] To connect (4) with the Pricing Equation (2), we assume $M_t R_{i,t} > 0$ and let $h = \ln(M_t R_{i,t})$ to obtain[10]

$$M_t R_{i,t} = 1 + \ln(M_t R_{i,t}) + z_{i,t}, \tag{5}$$

where the higher-order term of the expansion is

$$z_{i,t} \equiv \frac{1}{2} \times [\ln(M_t R_{i,t})]^2 \, e^{\lambda(\ln(M_t R_{i,t}))\ln(M_t R_{i,t})}.$$

It is important to stress that (5) is not an approximation but an exact relationship. Also, $z_{i,t} \geq 0$. Taking the conditional expectation of both sides of (5), using past information, denoted by $\mathbb{E}_{t-1}(\cdot)$, imposing the Pricing Equation, and rearranging terms, gives:

$$\begin{aligned} \mathbb{E}_{t-1}\{M_t R_{i,t}\} &= 1 + \mathbb{E}_{t-1}\{\ln(M_t R_{i,t})\} + \mathbb{E}_{t-1}(z_{i,t})\,, \text{ or,} && (6)\\ \mathbb{E}_{t-1}(z_{i,t}) &= -\mathbb{E}_{t-1}\{\ln(M_t R_{i,t})\}\,. && (7) \end{aligned}$$

Equation (7) shows that behavior of the conditional expectation of the higher-order term depends only on that of $\mathbb{E}_{t-1}\{\ln(M_t R_{i,t})\}$. Therefore, in general, it depends on lagged values of $\ln(M_t R_{i,t})$ and on powers of these lagged values. This will turn out to a major problem when estimating (1). To see it, denote by $\varepsilon_{i,t} = \ln(M_t R_{i,t}) - \mathbb{E}_{t-1}\{\ln(M_t R_{i,t})\}$ the innovation of $\ln(M_t R_{i,t})$. Let $\mathbf{R}_t \equiv (R_{1,t}, R_{2,t}, \ ..., R_{N,t})'$ and $\boldsymbol{\varepsilon}_t \equiv (\varepsilon_{1,t}, \varepsilon_{2,t}, \ ..., \varepsilon_{N,t})'$ stack respectively the returns $R_{i,t}$ and the forecast errors $\varepsilon_{i,t}$. From the definition of $\boldsymbol{\varepsilon}_t$ we have:

$$\ln(M_t \mathbf{R}_t) = \mathbb{E}_{t-1}\{\ln(M_t \mathbf{R}_t)\} + \boldsymbol{\varepsilon}_t. \tag{8}$$

Denoting $\mathbf{r}_t = \ln(\mathbf{R}_t)$, with elements $r_{i,t}$, and $m_t = \ln(M_t)$ in (8), and using (7) we get

$$m_t = -r_{i,t} - \mathbb{E}_{t-1}(z_{i,t}) + \varepsilon_{i,t}, \qquad \forall i. \tag{9}$$

Starting from (69), the covered, R^C, and the uncovered return, R^U, on foreign government bonds trade are, respectively,

$$R^C_{t+1} = \frac{{}_tF_{t+1}(1+i^*_{t+1})P_t}{S_tP_{t+1}} \qquad \text{and} \qquad R^U_{t+1} = \frac{S_{t+1}(1+i^*_{t+1})P_t}{S_tP_{t+1}}, \tag{10}$$

where ${}_tF_{t+1}$ and S_t are the forward and spot prices of foreign currency in terms of domestic currency, P_t is the dollar price level and i^*_{t+1} represents nominal net return on a foreign asset in terms of the foreign investor's preferences.

Using a forward version of (69) on both assets, and combining results, yields:

$$s_{t+1} - s_t = ({}_tf_{t+1} - s_t) - [\mathbb{E}_t(z_{U,t+1}) - \mathbb{E}_t(z_{C,t+1})] + \varepsilon_{U,t+1} - \varepsilon_{C,t+1}, \tag{11}$$

where the index i in $\mathbb{E}_t(z_{i,t+1})$ and $\varepsilon_{i,t+1}$ in (69) is substituted by either C or U, respectively for the covered and the uncovered return on trading foreign government bonds.

Under $\alpha_1 = 1$ and $\alpha_0 = 0$ in (1), taking into account (11), allows concluding that:

$$u_{t+1} = -[\mathbb{E}_t(z_{U,t+1}) - \mathbb{E}_t(z_{C,t+1})] + \varepsilon_{U,t+1} - \varepsilon_{C,t+1}.$$

Hence, by construction, the error term u_{t+1} is serially correlated because it is a function of current and lagged values of observables.[11] However, in most empirical studies, lagged observables are used as instruments to estimate (1) and test the null that $\alpha_1 = 1$, and $\alpha_0 = 0$. In that context, estimates of α_1 are biased and inconsistent, which may explain the finding that the average value of $\widehat{\alpha_1}$ is -0.88 for over 75 published estimates across various exchange rates and time periods. As far as we know, this is the first instance where FPP results are criticized in this fashion.

Because our goal is to relate the two puzzles, it is important to rephrase the FPP in the same language as the EPP. Recalling that rational expectations alone does not restrict the behavior of forward rates, since it is always possible to include a risk-premium term that would reconcile the time series behavior of the involved data, e.g., Fama (1984), the rejection of the null that $\alpha_1 = 0$, in favor of $\alpha_1 < 0$, only represents a true puzzle if reasonable risk measures cannot explain the empirical regularities of the data.

Here is where an asset-pricing approach may help, which is our starting point. 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.[12] The natural candidate for a theoretically sound model for pricing risk is the CCAPM of Lucas (1978) and Breeden (1979).

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

$$1 = \beta \mathbb{E}_t \left[\frac{u'(C_{t+1})}{u'(C_t)} R_{i,t+1} \right] \quad \forall i, \tag{12}$$

and, consequently,

$$0 = \mathbb{E}_t \left[\frac{u'(C_{t+1})}{u'(C_t)} \left(R_{i,t+1} - R_{j,t+1} \right) \right] \quad \forall i, j, \tag{13}$$

where $\beta \in (0,1)$ is the discount factor in the representative agent's utility function, $R_{i,t+1}$ and $R_{j,t+1}$ are, respectively, the real gross return on assets i and j at time $t+1$ and, C_t is aggregate consumption at time t. In other words, under the CCAPM, $M_{t+1} = \beta u'(C_{t+1})/u'(C_t)$.

Let the standing representative agent be a U.S. investor who can freely trade domestic and foreign assets.[13] Define the covered and the uncovered return on trading foreign government bonds as in (10) and substitute R^C for R_i and R^U for R_j in (12) to get

$$0 = \mathbb{E}_t \left[\frac{u'(C_{t+1})}{u'(C_t)} \frac{P_t(1+i^*_{t+1})[{}_tF_{t+1} - S_{t+1}]}{S_t P_{t+1}} \right]. \tag{14}$$

Assuming that preferences exhibit constant relative risk aversion, Mark (1985) estimated the parameter α in $u(C) = C^{1-\alpha}(1-\alpha)^{-1}$, applying Hansen's (1982) Generalized Method of Moments (GMM) to (14), reporting an estimated coefficient of relative risk aversion, $\widehat{\alpha}$, above 40. He then tested the over-identifying restrictions to assess the validity of the model, rejecting them when the forward premium and its lags are used as instruments. Similar results were reported later by Modjtahedi (1991). Using a different, larger data set, Hodrick (1989) reported estimated values of $\widehat{\alpha}$ above 60, but did not reject the over-identifying restrictions, while Engel (1996) reported some estimated $\widehat{\alpha}$'s in excess of 100. A more recent attempt to use euler equations to account for the FPP is Lustig and Verdelhan (2006 a), where risk aversion in excess of 100 is needed to price the forward premium on portfolios of foreign currency.

Are we to be surprised with these findings? If we recall that the EPP is identified with the failure of consumption-based kernels to explain the excess return of equity over risk-free short term bonds – $R_{i,t+1} = (1+i^{SP}_{t+1})P_t/P_{t+1}$ and $R_{j,t+1} = (1+i^{b}_{t+1})P_t/P_{t+1}$, in (13), where i^{SP}_{t+1} is nominal return on S&P500 and i^{b}_{t+1} nominal return on the U.S. Treasury Bill – with reasonable parameters of risk aversion for (13),[14] why should we expect these same consumption-based models not to generate the FPP? Indeed, we should expect the opposite.

The inexistence of a widely accepted model to account for risk is partly to blame for the separation of the research agendas involving the two puzzles. There is, however, an additional reason. Another characteristic of the FPP may have played a role in the separation of these two research agendas: the predictability of returns on currency speculation. Because $\widehat{\alpha_1} < 0$ and significant, given that the auto-correlation of risk premium is very persistent, interest-rate differentials predict excess returns. Although predictability in equity markets has by now been extensively documented, it was *not* viewed as a defining feature of the EPP, back then. It was, however, a defining feature of the FPP, which has lead Engel (1996, p. 155), for example, to write: "International economists face not only the problem that a high degree of risk aversion is needed to account for estimated values of [the risk premium demanded by a rational agent]. There is also the question of why the forward premium is such a good predictor of $s_{t+1} -_t f_{t+1}$. There is no evidence that the proposed solutions to the puzzles in domestic financial markets can shed light on this problem."

Predictability is now acknowledged to be present in domestic markets as well, in the context of the equity premium. Dividend-to-price ratio, and other variables are capable of predicting returns, which means that, once again, we should be suspicious that the same underlying forces may account for asset behavior in both markets.

Before describing our strategy it is important to draw attention to the fact that pricing excess returns is crucial, but should not be the sole goal of asset-pricing theory. Returns, and not only excess returns need to be priced, and accomplishing both is a much harder task. To make the point as stark as possible, let us get back to (12). When we substitute R^C for R_i and R^U for R_j, we get

$$1 = \mathbb{E}_t \left[\beta \frac{u'(C_{t+1})}{u'(C_t)} \frac{{}_tF_{t+1}(1+i^*_{t+1})P_t}{S_t P_{t+1}} \right] \quad \text{and} \quad 1 = \mathbb{E}_t \left[\beta \frac{u'(C_{t+1})}{u'(C_t)} \frac{S_{t+1}(1+i^*_{t+1})P_t}{S_t P_{t+1}} \right]. \tag{15}$$

It turns out that in the canonical model, e.g., Hansen and Singleton (1982, 1983, 1984), the parameter of risk aversion is the inverse of the intertemporal elasticity of substitution, meaning that if one wants to accept a high risk aversion, one generates implausibly high and volatile interest rates.

Accordingly, if one wants to identify the structural parameter β, in an econometric sense, one cannot resort to direct estimation of excess returns (e.g., 14), but rather to joint estimation of the two euler equations for returns (e.g., 15), or to any linear rotation of them. It is, therefore, important to make a distinction between studies that test the over-identifying restrictions jointly implied by returns and those that test the ones implied by excess returns. For the latter no-rejection may be consistent with *any* value for β, including inadmissible ones.[15]

In our view, a successful consumption-based model must account for asset prices everywhere (domestically and abroad), as well as price returns, excess returns, and many new facts recently evidenced in the extensive empirical research that has been in a great deal sparked by the theoretical developments of the late seventies - see, for example, Cochrane (2006).

3. Our Strategy

The stated purpose of this paper is to relate the EPP and the FPP. The fact that, as of this moment, no satisfactory consumption-based model derived from the primitives of the economy can account for asset behavior in either market is a hint that it is generating the two puzzles. We argue here in favor of an indirect approach. We do not need a proper consumption model for the pricing kernel to link the two puzzles. All we need is a strategy to extract a proper pricing kernel from return data, showing that it prices both the domestic and the foreign-exchange returns and excess returns. This isolates current consumption-based kernels as the most likely culprits for mispricing these two markets. Of course, a final proof that these puzzles are linked in this fashion can only be obtained when we finally have a proper consumption-based model to price assets. That will explain why current models fail, something we cannot do here.

Following Harrison and Kreps (1979) Hansen and Richard (1987), and Hansen and Jagganathan (1991), we write the system of asset-pricing equations,

$$1 = \mathbb{E}_t \left[M_{t+1} R_{i,t+1} \right], \quad \forall i = 1, 2, \cdots, N, \tag{16}$$

leading to

$$0 = \mathbb{E}_t \left[M_{t+1} \left(R_{i,t+1} - R_{j,t+1} \right) \right], \quad \forall i, j. \tag{17}$$

We combine statistical methods with these Asset Pricing Equations to devise pricing-kernel estimates as projections of SDF's on the space of returns, i.e., the SDF mimicking portfolio, which is unique even under incomplete markets. We denote the latter by M^*_{t+1}. These pricing kernels do not depend on any assumptions about a functional form for preferences, but solely on returns. In this sense, these methods are *preference free*, despite the fact that they have a strong footing on theory as a consequence of the use of the Asset Pricing Equation.

Our exercise consists in exploring a large cross-section of U.S. time-series stock returns to construct return-based pricing kernel estimates satisfying the Pricing Equation (16) for that group of assets. Then, we take these SDF estimates and use them to price assets not used in constructing them. Therefore, we perform a genuine out-of-sample forecasting exercise using SDF mimicking portfolio estimates, avoiding in-sample over-fitting.

We cannot overstress the importance of out-of-sample forecasting for our purposes. Our main point in this paper is to show that the forward- and the equity-premium puzzle are intertwined. Under the law of one price, an SDF exists that prices all assets, necessarily. Thus, an in-sample exercise would only provide evidence that the forward-premium puzzle is not simply a consequence of violations of the law of one price. We aim at showing more: a SDF can be constructed using only domestic assets, i.e., using the same source of information that guides research regarding the equity premium puzzle, and still price foreign assets. It is our view that this SDF is to capture the growth of the marginal utility of consumption in a model yet to be written.[16]

The main rationale for our methodological choice is the fact that, however successful, a consumption-based model will not perform better in pricing tests than its related mimicking portfolio. This is the main trust of Hansen and Jagganathan (1991). The mimicking portfolio, thus, represents an upper bound on the pricing capability of any model. Given our purposes an alternative approach would be to try and relate the EPP and the FPP using a volatile consumption kernel constructed with high aversion values. The trouble with this approach is that the work by Hansen and Singleton (1982, 1983, 1984) generated a consensus that the over-identifying restrictions of traditional consumption models are often rejected when used to price not only excess returns but also returns, i.e., when both the discount factor β and the risk-aversion coefficient γ are identified in an econometric sense in canonical models. We show that this pattern is present in our data set as well.

Before moving on to the description of our methodology it is worth mentioning that, in dispensing with consumption data, our paper parallels those of Hansen and Hodrick (1983), Hodrick and Srivastava (1984), Cumby (1988), Huang (1989), and Lewis (1990), all of which implemented latent variable models that avoid the need for specifying a model for the pricing kernel by treating the return on a benchmark portfolio as a latent variable. [17] Also related 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 do not perform any out-of-sample exercises and do not try to relate the two puzzles.

Finally, Backus et al. (1995) ask whether a pricing kernel can be found that satisfies, at the same time, log-linearized versions of

$$0 = \mathbb{E}_t \left[M_{t+1}^* \frac{P_t(1+i_{t+1}^*)[{}_tF_{t+1} - S_{t+1}]}{S_t P_{t+1}} \right] \text{ and,} \tag{18}$$

$$R_{t+1}^f = \frac{1}{\mathbb{E}_t(M_{t+1}^*)}, \tag{19}$$

where R_{t+1}^f is the risk-free rate of return. The nature of the question we implicitly answer 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 U.S.[18]

3.1. Econometric Tests

Assume that we are able to approximate well enough a time series for the pricing kernel, M_{t+1}^*. Next, we show how to use this approximation to implement direct pricing tests for the forward and the equity-premium, in an euler equation framework. In Section we discuss how to construct this time series for M_{t+1}^* using asset-return information.

3.1.1. Pricing Test

In the context of the SDF mimicking portfolio, euler equations (16) and (17) must hold for all assets and portfolios. If we had observations on M_t^*, then we would only need return data to test directly whether they held. Of course, M_t^* is a latent variable. Despite that, if we had a consistent estimator for M_t^* based on return data, and a large enough sample, so that M_t^* and their estimators are "close enough," we could still directly test the validity of these euler equations using return data alone: the estimators of M_t^* are a function of return data and returns are also used to verify the Asset Pricing Equation. In this case, we do not have to perform log-linear approximations of these euler equations, nor do we have to impose the stringent restrictions that returns are log-Normal and Homoskedastic to test theory. Because we want our tests to be *out-of-sample* the returns used to construct the estimates of M_t^* will not be the ones used directly in the euler equations when testing theory.

Consider z_t to be a vector of instrumental variables, which are all observed up to time t, therefore measurable with respect to $\mathbb{E}_t(\cdot)$. Employing *scaled returns* and *scaled excess-returns* – defined as $R_{i,t+1} \times z_t$ and $(R_{i,t+1} - R_{j,t+1}) \times z_t$, respectively – we are able to test the conditional moment restrictions associated with the euler equations and consequently to derive the implications from the presence of information. This is particularly important for the FPP, since, when the CCAPM is employed, the over-identifying restriction associated with having the own current forward premium as an instrument is usually rejected: a manifestation of its predictive power.

Multiply

$$0 = \mathbb{E}_t \left[M_{t+1}^* \frac{P_t(1+i_{t+1}^*)[{}_tF_{t+1} - S_{t+1}]}{S_t P_{t+1}} \right], \text{ and,} \tag{20}$$

$$1 = \mathbb{E}_t \left[M_{t+1}^* \frac{S_{t+1}(1+i_{t+1}^*)P_t}{S_t P_{t+1}} \right] \tag{21}$$

by z_t and apply the Law-of-Iterated Expectations to get, respectively,

$$0 = \mathbb{E} \left\{ \left[M_{t+1}^* \frac{P_t(1+i_{t+1}^*)[{}_tF_{t+1} - S_{t+1}]}{S_t P_{t+1}} \right] \times z_t \right\}, \text{ and,} \tag{22}$$

$$0 = \mathbb{E} \left\{ \left[M_{t+1}^* \frac{S_{t+1}(1+i_{t+1}^*)P_t}{S_t P_{t+1}} - 1 \right] \times z_t \right\} \tag{23}$$

Equations (22) and (23) form a system of orthogonality restrictions that can be used to assess the pricing behavior of estimates of M_{t+1}^* with respect to the components of the forward premium or any linear rotation of them. Equations in the system can be tested separately or jointly. In testing, we employ a generalized method-of-moment (GMM) perspective, using (22) and (23) as a natural moment restriction to be obeyed. Consider parameters μ_1 and μ_2 in:

$$0 = \mathbb{E}\left\{\left[M^*_{t+1}\frac{P_t(1+i^*_{t+1})[_tF_{t+1}-S_{t+1}]}{S_tP_{t+1}}-\mu_1\right]\times z_t\right\}, \text{ and,} \quad (24)$$

$$0 = \mathbb{E}\left\{\left[M^*_{t+1}\frac{S_{t+1}(1+i^*_{t+1})P_t}{S_tP_{t+1}}-(1-\mu_2)\right]\times z_t\right\} \quad (25)$$

We assume that there are enough elements in the vector z_t for μ_1 and μ_2 to be over identified. In order for (22) and (23) to hold, we must have $\mu_1 = 0$ and $\mu_2 = 0$, and the over-identifying restriction $T \times J$ test in Hansen (1982) should not reject them. This constitutes the econometric testing procedure implemented in this paper to examine whether the FPP holds when return-based pricing kernels are used.

A similar procedure can be implemented for the domestic market equations, in order to investigate domestic stylized facts that escape consumption based models. To analyze the EPP, the system of conditional moment restrictions is given by:

$$0 = \mathbb{E}\left\{\left[M^*_{t+1}\frac{(i^{SP}_{t+1}-i^b_{t+1})P_t}{P_{t+1}}-\mu_1\right]\times z_t\right\}, \text{ and} \quad (26)$$

$$0 = \mathbb{E}\left\{\left[M^*_{t+1}\frac{(1+i^{SP}_{t+1})P_t}{P_{t+1}}-(1-\mu_2)\right]\times z_t\right\}, \quad (27)$$

where i^{SP}_{t+1} and i^b_{t+1} are respectively the returns on the S&P500 and on a U.S. government short-term bond, and we also test whether $\mu_1 = 0$ and $\mu_2 = 0$, and check the appropriateness of the over-identifying restrictions using Hansen's $T \times J$ test[19].

Beyond the high equity Sharp ratio or the reported power of the dividend-price ratio to forecast stock-market returns, the pattern of cross-sectional returns of assets exhibit some "puzzling aspects" as the "size" and the "value" effects – e.g. Fama and French (1996) and Cochrane (2006) –, i.e., the fact that small stocks and of stocks with low market values relative to book values tend to have higher average returns than other stocks.

We follow Fama and French (1993) in using our pricing kernels to try and account for their stock-market factors; zero-cost portfolios which are able to summarize these effects, explaining average returns on stocks and bonds. In this case, the system of conditional moment restrictions is given by:

$$0 = \mathbb{E}\left\{\left[M_{t+1}^{*}(R_m - R_f)_{t+1} - \mu_1\right] \times z_t\right\}, \tag{28}$$

$$0 = \mathbb{E}\left\{\left[M_{t+1}^{*}HML_{t+1} - \mu_2\right] \times z_t\right\} \quad \text{and} \tag{29}$$

$$0 = \mathbb{E}\left\{\left[M_{t+1}^{*}SMB_{t+1} - \mu_3\right] \times z_t\right\}, \tag{30}$$

where $Rm - Rf$, the excess return on the market, is the value-weighted return on all NYSE, AMEX, and NASDAQ stocks minus the one-month Treasury bill rate, HML (High Minus Low) is the average return on two value portfolios minus the average return on two growth portfolios and SMB (Small Minus Big) is the average return on three small portfolios minus the average return on three big portfolios.[20] We, again, test $\mu_1 = 0$, $\mu_2 = 0$ and $\mu_3 = 0$, and check the appropriateness of the over-identifying restrictions using Hansen's $T \times J$ test.

It worth recalling that we employ an euler-equation framework, something that was missing in the forward-premium literature after Mark (1985). Since the two puzzles are manifest in logs and in levels, by working directly with the Pricing Equation we avoid imposing stringent auxiliary restrictions in hypothesis testing, while keeping the possibility of testing the conditional moments through the use of lagged instruments along the lines of Hansen and Singleton (1982 and 1984) and Mark. We hope to have convinced our readers that the log-linearization of the euler equation is an unnecessary and dangerous detour. Any criticism arising from the use of the log-linear approximation is avoided here.

An important feature of our testing procedure is its out-of-sample character. To preserve the temporal structure of the euler equations, we perform out-of-sample tests in the cross-sectional dimension, i.e., the returns used in estimating M_{t+1}^{*} exclude the return of the assets appearing directly in our main tests. Therefore, there is no reason for the Asset Pricing Equation to hold for the assets not used in estimating M_{t+1}^{*}.

3.1.2. Instruments

There seems to be a consensus in the return forecasting literature about the rejection of the time-invariant excess returns hypothesis. However, the question of which variables can be considered as good predictors for returns is still open. Hence, the choice of a representative set of forecasting instruments plays certainly an important role and highlights the relevance of the conditional tests.

Taking into account the fact that expected returns and business cycles are correlated, as documented in Fama and French (1989), for both domestic and international markets, we use the following macroeconomic variables: real consumption instantaneous growth rates, real GDP instantaneous growth rates, and the consumption-GDP ratio. However, since in our exercise, the forecasting variables, the pricing kernels, and the excess returns, are all based on market prices, we also include specific financial variables as instruments, carefully choosing them based on their forecasting potential.

Regarding the FPP, besides using as instruments the past values for the covered and uncovered returns on trading of the respective foreign government bond, we also use the current value of the forward premium, since the well documented predictability power of this variable is a defining feature of this puzzle.

For the Fama and French (1993) portfolios and the EPP, we use lagged values of the returns on relevant assets as instruments, since one should not omit the possibility that returns could be predictable from past returns for any financial market. Finally, for the EPP we still use the dividend-price ratio, following Campbell and Shiller (1988) and Fama and French (1988), who show evidences of the good performance of this variable as a predictor of stock-market returns.

3.2. Return-Based Pricing Kernels and the SDF Mimicking Portfolio

The basic idea behind estimating return-based pricing kernels with asymptotic techniques is that asset prices (or returns) convey information about the intertemporal marginal rate of substitution in consumption. If the Asset Pricing Equation holds, all returns must have a common factor that can be removed by subtracting any two returns. A common factor is the SDF mimicking portfolio M^*_{t+1}. Because every asset return contains "a piece" of M^*_{t+1}, if we combine a large enough number of returns, the average idiosyncratic component of returns will vanish in limit. Then, if we choose our weights properly, we may end up with the common component of returns, i.e., the SDF mimicking portfolio.

Although the existence of a strictly positive SDF can be proved under no arbitrage, uniqueness of the SDF is harder to obtain, since under incomplete markets there is, in general, a continuum of SDF's pricing all traded securities. However, each M_{t+1} can be written as $M_{t+1} = M^*_{t+1} + \nu_{t+1}$ for some ν_{t+1} obeying $\mathbb{E}_t [\nu_{t+1} R_{i,t+1}] = 0 \ \forall i$. Since the economic environment we deal with is that of incomplete markets, it only makes sense to devise econometric techniques to estimate the unique SDF mimicking portfolio – M^*_{t+1}.

There are two basic techniques employed here to estimate M^*_{t+1}. The first one uses principal-component and factor analyses. It can be traced back to the work of Ross (1976), developed further by Chamberlain and Rothschild (1983), and Connor and Korajczyk (1986, 1993). A recent additional reference is Bai (2005). This method is asymptotic: either $N \to \infty$ or $N, T \to \infty$, relying on weak law-of-large-numbers to provide consistent estimators of the SDF mimicking portfolio – the unique systematic portion of asset returns. An alternative to this asymptotic method is to use a method-of-moment approach, constructing algebraically the unique projection of any SDF on the space of returns. This can be achieved through the following linear combination of traded returns:

$$M^*_{t+1} \equiv \mathbf{1}' \left[\mathbb{E}(\mathbf{R}_{t+1}\mathbf{R}'_{t+1})\right]^{-1} \mathbf{R}_{t+1},$$

where $\mathbf{1}$ and $\mathbf{R}_{t+1}$ are $N \times 1$ vectors of ones and of traded returns respectively. This technique was proposed by Hansen and Jagganathan (1991) to estimate the SDF mimicking portfolio using the Pricing Equation. For sake of completeness, we present a summary account of these the first method in section , as well as a more complete description of both of them in the Appendix.

3.2.1. Multifactor Models

Factor models summarize the systematic variation of the N elements of the vector $\mathbf{R}_t = (R_{1,t}, R_{2,t}, \ ..., R_{N,t})'$ using a reduced number of K factors, $K < N$. Consider a K-factor model in $R_{i,t}$:

$$R_{i,t} = a_i + \sum_{k=1}^{K} \beta_{i,k} f_{k,t} + \eta_{it}, \tag{31}$$

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

$$\operatorname*{plim}_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \eta_{i,t} = 0.$$

Denote by $\Sigma_{\mathbf{r}} = \mathbb{E}(\mathbf{R}_t\mathbf{R}'_t) - \mathbb{E}(\mathbf{R}_t)\,\mathbb{E}(\mathbf{R}'_t)$ the variance-covariance matrix of returns. The first principal component of the elements of $\mathbf{R}_t$ is a linear combination $\theta'\mathbf{R}_t$ with maximal variance subject to the normalization that θ has unit norm, i.e., $\theta'\theta = 1$. 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 $\mathbf{R}_t$ are consistent estimates of the $f_{k,t}$'s. Factor loadings can be estimated consistently by simple OLS regressions of the form (31).

It is straightforward to connect principal-component and factor analyses with the Pricing Equation, delivering a consistent estimator for M_t^*. Given estimates of a_i, $\beta_{i,k}$, and $f_{k,t}$ in (31), one can write their respective expected-beta return expression:

$$\mathbb{E}(R_i) = \gamma + \sum_{k=1}^{K} \beta_{i,k}\lambda_k, i = 1, 2, ..., N$$

where λ_k is interpreted as the price of the k-th risk factor. 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 λ coefficients directly by

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

and consequently to estimate only γ via a cross-sectional regression[21]. Given these coefficients, one can easily get an estimate of M_t^*,

$$\widetilde{M_t^*} \equiv a + \sum_{k=1}^{K} b_k f_{k,t}$$

where (a, b) is related to (λ, γ) through

$$a \equiv \frac{1}{\gamma} \text{ and } b \equiv -\gamma \left[cov(ff')\right]^{-1} \lambda,$$

It is easy then to see the equivalence between the beta pricing model and the linear model for the SDF. More, it is immediate that

$$\mathbb{E}(\widetilde{M_t^*} R_{i,t}) = 1, \quad i = 1, 2, ..., N.$$

The number of factors used in the empirical analysis is an important issue. We expect K to be rather small. We followed Lehmann and Modest (1988) and Connor and Korajczyk (1988), taking the pragmatic view whereby increasing K until the estimate of M_t^* changed very little due to the last increment in the number of factors.

4. Empirical Results

4.1. Data and Summary Statistics

In principle, whenever econometric or statistical tests are performed, it is preferable to employ a large data set either in the time-series (T) or in the cross-sectional dimension (N). Regarding the FPP, the main limitation is the fact that the Chicago Mercantile Exchange, the pioneer of the financial-futures market, only launched currency futures in 1972. In addition to that, only futures data for a few developed countries are available since then. In order to have a common sample for the largest set of countries possible, we considered here U.S. foreign-exchange data for Canada, Germany, Japan, Switzerland and the U.K., covering the period from 1990:1 to 2004:3, on a quarterly frequency. In order to extend the time span of used here, keeping a common sample for all countries, we would have to accept a drastic reduction in the number of countries, which we regard as an inferior choice.

Spot and forward exchange-rate returns were transformed into U.S.$ real returns using the consumer price index in the U.S. 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. To study the EPP we used the U.S.$ real returns on the S&P500 and on 90-day T-Bill. Real returns were obtained using the consumer price index in the U.S.

A second ingredient for testing these two puzzles is to estimate return-based pricing kernels. Again, in choosing return data, we had to deal with the trade-off between N and T. In order to get a larger N, one must accept a reduction in T: disaggregated returns are only available for smaller time spans than aggregated returns. The database used here to estimate the SDF is comprised of U.S.$ real returns on two hundred U.S. stocks – those with the 200 largest volumes according to CRSP database. Therefore, it is completely U.S. based and available at a very disaggregated level. Our choice of returns to estimate the SDF mimicking portfolio is a direct response to Cochrane's (2001) criticism of in-sample over-fitting: the return data used to construct SDF estimates is not the same used to construct excess returns in foreign markets. Hence, our pricing tests are out-of-sample in the cross-sectional dimension (assets).

All macroeconomic variables used in econometric tests were extracted from FED's FRED database. We also employed additional forecasting financial variables that are specific to each test performed, and are listed in the appropriate tables of results. The Fama-French benchmark factors series were extracted from the French data library. In terms of the notation used in the tables below, we adopted the following: the estimate of M_t^* using multi-factor models is labelled $\widetilde{M_t^*}$, while that using the projection in Hansen and Jagganathan (1991) is labelled $\overline{M_t^*}$.

Table I presents a summary statistic of our database over the period 1990:1 to 2004:3. The average real return on the covered trading of foreign government bonds range from 0.91% to 1.78% a year, while that of uncovered trading range from 1.69% to 4.91%. The real return on the S&P500 is 8.78% at an annual rate, while that of the 90-day T-Bill is 1.42%, with a resulting excess return of 7.28%. As expected, real stock returns are much more volatile than the U.S. Treasury Bill return – annualized standard deviations of 16.82% and 1.10% 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 1.69% to 4.91%, while standard deviations range from 5.52% to 15.50%.

We computed the Sharpe ratio for the U.S. stock market to be 0.44, while the Sharpe ratio of the uncovered trading of foreign bonds ranges from -0.01 to 0.38. 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 reason behind the EPP. Since the higher the Sharpe ratio, the tighter the lower bound on the volatility of the pricing kernel, a natural question that arises is the following: may we regard this fact as evidence that a kernel that prices correctly the equity premium would also price correctly the forward premium? We will try to answer this question here in an indirect way.

4.2. SDF Estimates

In constructing $\widetilde{M_t^*}$, from the real returns on the two hundred most traded (volume) U.S. stocks, we must first choose the number of factors, i.e., how many pervasive factors are needed to explain reasonably well the variation of these 200 stock returns? Following Lehmann and Modest (1988), Connor and Korajczyk (1988), and most of the empirical literature, we took a pragmatic view, increasing the number of factors until $\widetilde{M_t^*}$ changed very little with respect to choosing an additional factor. We concluded that 6 factors are needed to account for the variation of our 200 stock returns: starting from 3 factors, increasing this number up to 6, implies very different estimates of M_t^*. However, starting from 6 factors, increasing this number up to 8, implies practically the same estimate of M_t^*. Our choice (6 factors) is identical to that of Connor and Korajczyk (1993), who examined returns from stocks listed on the New York Stock Exchange and the American Stock Exchange.

Looking at the final linear combination of returns that comprise $\widetilde{M_t^*}$, we list the following most relevant stocks in its composition: 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).

In constructing $\overline{M_t^*}$, a practical numerical problem had to be faced, which is how to invert the second-moment matrix $\mathbb{E}\left(\mathbf{R}_{t+1}\mathbf{R}'_{t+1}\right)$ – a square matrix of order 200. Standard inversion algorithms broke down and we had to resort to the Moore-Penrose generalized inverse technique.

The estimates of M_t^* – $\overline{M_t^*}$ and $\widetilde{M_t^*}$ – are plotted in Figure 1, which also includes their summary statistics. Their means are slightly below unity, 0.962 and 0.977 respectively. Moreover, the mimicking portfolio estimate is about twice more volatile than the multi-factor model estimate. The correlation coefficient between $\overline{M_t^*}$ and $\widetilde{M_t^*}$ is 0.407.

4.3. Pricing-Test Results

Table II presents results of the over-identifying-restriction tests when consumption-based kernels are employed and excess returns are represented by the equity premium in the U.S. These results will be later compared to those using return-based kernels. We considered three types of preference representations here: standard CRRA, following Hansen and Singleton (1982, 1983, 1984), Kreps-Porteus, following Epstein and Zin (1991), and External Habit, following Abel (1990). Tests are conducted separately for the euler equation for excess returns and for the two euler equations for returns. In the former case, the discount rate β is not identified, which is not a feature of the latter.

The top portion of Table II presents test results when excess returns (U.S. equity versus U.S. government bonds) are considered, i.e., when β is not identified. In this case, the over-identifying-restriction test does not reject the null at 5% significance regardless of the type of the utility function we considered. Since β is not identified, the euler equation for excess returns is consistent with any arbitrary value of β. Notice that estimates of the constant relative-risk aversion coefficient are in excess of 160 for all preference specifications used, which is similar to the results obtained by Lustig and Verdelhan (2006 a).[22]

The lower portion of Table II presents test results for a system of two euler equations for U.S. equities and government bonds. Here, both the discount rate β and the risk-aversion coefficient α are identified. A completely different result with respect to Table II emerges in this case: with very high confidence, the over-identifying-restriction test rejects the null regardless of the preference-specification being considered; estimates of α are relatively small, but significant; and β estimates are close to unity and significant as well. Because of the overwhelming rejection of the over-identifying-restriction test, we conclude that the EPP is a feature of our data set: we cannot reconcile data and theory using standard econometric tests and at the same time obtain "reasonable" parameter estimates. When testing did not reject the over-identifying-restrictions – results in Table II – β was not identified, and estimates of α were in excess of 160, which are far from what we may call reasonable.

Table III presents single-equation equity-premium test results when return-based pricing kernel estimates are used in place of consumption-based kernels. A variety of macroeconomic and financial instruments, including up to their own two lags, are employed in testing. At the 5% significance level, when $\widetilde{M_t^*}$ is used, there is only one instance when the $T \times J$ statistic rejects the null. This happens when the dividend-price ratio $\frac{D_t}{P_t}$ (up to its own two lags) is used as an instrument. When the mimicking portfolio is estimated using $\overline{M_t^*}$, there is no rejection of the $T \times J$ statistic at the 5% significance level. Also, in all instances, estimates of $\mu_1$and μ_2 obey $\mu_1 = 0$ and $\mu_2 = 0$ at the usual confidence levels.

Table IV presents the first set of results regarding the forward-premium puzzle, in a single-equation context, where $\widetilde{M_t^*}$ and $\overline{M_t^*}$, as estimators of M_t^*, are used to price the excess return of uncovered over covered trading with foreign government bonds, while

Table V presents tests, where $\widetilde{M_t^*}$ and $\overline{M_t^*}$ are used to price the return on uncovered trading with foreign government bonds. In both tables, at the 5% significance level, there is not one single rejection either in mean tests for $\mu_1$and μ_2 or for the $T \times J$ statistic.

Table VI presents equity-premium tests for systems, when $\widetilde{M_t^*}$ and $\overline{M_t^*}$ are used. No theoretical restriction is rejected here either by individual tests for $\mu_1$and μ_2 or for the $T \times J$ statistic. Table VII presents system tests for the Fama-French portfolios. We do not reject the null for the market excess return $(R_m - R_f)$ and for the small/big (SMB) portfolio, but not for the high/low (HML) one. It is worth stressing that, even then, the over-identifying-restriction test does not reject the null.

Finally, Table VIII presents system tests for the FPP when $\widetilde{M_t^*}$ and $\overline{M_t^*}$ are used respectively. At 5% significance, there is a single rejection (out of 10) of the over-identifying-restriction test for German bonds with $\overline{M_t^*}$. Also, for British bonds, there is evidence that $\mu_1 \neq 0$, although the $T \times J$ statistic does not reject the null.

4.4. Discussion

In this paper, we first questioned the standard testing procedure of the FPP, relying on estimates $\widehat{\alpha_0}$ and $\widehat{\alpha_1}$, obtained from running

$$s_{t+1} - s_t = \alpha_0 + \alpha_1({}_tf_{t+1} - s_t) + u_{t+1},$$

from a theoretical point of view. Our key point is that ${}_tf_{t+1} - s_t$ and u_{t+1} are correlated, and that lagged observables are not valid instruments. Since these are exactly the instrumental variables used in obtaining $\widehat{\alpha_0}$ and $\widehat{\alpha_1}$, tests of the FPP relying on such estimates are biased and inconsistent.

Next, we show evidence of the EPP in our data, if one takes the EPP to mean "the failure of consumption-based kernels to explain the excess return of equity over risk-free short term bonds with reasonable parameters values for risk aversion." This is a consequence of the results obtained in Table II, where system tests involving the returns of these two assets overwhelmingly rejected the implied over-identifying restrictions, and single-equation estimates of the risk-aversion coefficient were in excess of 160, regardless of the preference specification employed. A risk-aversion coefficient greater than 160 cannot be called "reasonable" under any circumstance, especially if it is obtained only in models where the discount rate coefficient β is not identified in the econometric sense.

Recent work by Lustig and Verdelhan (2006a) estimates the coefficient of relative risk aversion to fit an euler equation under more general preferences than those used in Mark (1985) and Hodrick (1987) capable of pricing the returns on eight different portfolios of foreign currencies. Because they do not try to price returns on these portfolios, one cannot fully assess the adequacy of the pricing kernels they generate.

Next, we showed that, using return-based pricing kernels, in an euler-equation setup, we are able to properly price returns and excess returns of assets that comprise the equity premium and the forward premium puzzles; see results presented in Tables 3 through 8, where several econometric tests were performed, either in single equations or in systems, and using two distinct estimates of $M_t^* - \overline{M_t^*}$ and $\widetilde{M_t^*}$. These test results are very informative for at least two reasons. First, even if we do not have a consumption model that delivers a proper pricing kernel, the mimicking portfolio is a valid kernel. Second, our tests have an out-of-sample character in the cross-sectional dimension.

One important element of our testing procedure is that, when the ratio ${}_tF_{t+1}/S_t$ is used as an instrument, the theoretical restrictions tested were not rejected, leading to the conclusion that ${}_tF_{t+1}/S_t$ has no predictive power for $M_{t+1}^* \frac{P_t(1+i_{t+1}^*)}{P_{t+1}} \frac{{}_tF_{t+1}-S_{t+1}}{S_t}$. Hence, although the excess returns on uncovered over covered trading with foreign bonds are predictable, "risk adjusted" excess returns are not. This raises the question that predictability results using (1) may just be an artifact of the log-linear approximation of the euler equation for excess returns. This is a very important result, since predictability of $s_{t+1} - s_t$ in (1) is a defining feature of the forward-premium puzzle.

Given that Mark (1985) found evidence of the FPP in proper econometric tests, also found by Hansen and Singleton (1982, 1984) regarding the EPP[23], and that we show enough evidence that proper estimates of M_t^* do not misprice returns comprising the equity premium and the forward premium, we conclude that the EPP and the FPP are "two symptoms of the same illness" – the poor (although steadily improving) performance of current consumption-based pricing kernels to price asset returns or excess returns. Our result takes the correlation with the pricing kernel as the appropriate measure of risk, and adds to the body of evidence that "explains" the forward premium as a risk premium. The reason why we quote the term explains is because, we have not tried to advance on the explanation of either puzzle, but simply to show that the two are related.

Even though we believe that we have elements to give a positive answer to the question posed in the title of this paper, a different question we may ask is: should the forward premium be regarded as a reward for risk taking? If we take the covariance with M^*_{t+1} as the relevant measure of risk, then our answer is yes. In this sense we side with the position 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. However, this is not without controversy. Citing Engel (1996, 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."

5. Conclusion

Previous research has cast doubt on whether consumption-based pricing kernels were capable of correctly pricing the equity and the forward premium,[24] generating respectively the EPP and the FPP, with two different literatures. Here, we propose a fresh look into the relationship between these well known puzzles. We employ an asset-pricing approach. Our starting point is the Asset Pricing Equation, coupled with the use of consistent estimators of the SDF mimicking portfolio. They are a function of return data alone and do not depend on aggregate consumption or on any parametric representation for preferences. In this context, we first show that, our estimated return-based kernels price correctly the equity and the forward premium, as well as the individual returns that compose them. Our estimates are constructed using domestic (U.S.) returns alone.

Based on our empirical results, we go one step further and ask whether the EPP and the FPP are but two symptoms of the same illness – the inability of standard (and augmented) consumption-based pricing kernels to price asset returns or excess-returns. Given the tendency in the profession of generating new research agendas whenever a new empirical regularity that cannot be accounted for by our models is discovered, it is important to always ask whether these are distinct phenomena or if they are but two manifestations of a same problem. Otherwise, in the limit, we could find as many puzzles as there were assets. Since the number of assets in any real economy is large, would it make any sense to investigate all of them separately? Obviously not. What we are able to show is that, indeed, regarding the EPP and the FPP, finding a model that does account for either puzzle is bound to double its prize by accounting for the other.

Our empirical tests are robust to important sources of misspecification incurred by the previous literature: inappropriate log-linear approximation of the euler equation for returns and inappropriate models for consumption-based kernels.[25]

Our return-based pricing kernels, constructed using only U.S.-stock returns, are orthogonal to past information that is usually known to forecast undiscounted excess returns. Moreover, they are also able to price the equity premium for the U.S. and the exchange-rate forward premium for five distinct developed economies: United Kingdom, Canada, Germany, Japan and Switzerland. In our tests, we found that the *ex-ante* forward premium is not a predictor of discounted excess returns, as is usually found when a log-linear approximation of the Asset Pricing Equation is employed in testing. This evidence, coupled with our theoretical discussion on the log-linear approximation of the euler equation, cast doubt on the predictability of exchange-rate changes. In our opinion predictability may be a consequence of the inappropriateness of the log-linear approximations previously employed in testing theory.

In our tests, although consumption-based kernels are not able to correctly price returns, return-based kernels are. This provides the basis for believing that the two puzzles are two symptoms of the same illness, being therefore more profitable to concentrate efforts on a single research agenda, focused on rethinking consumption-based pricing kernels. As stressed before, the final proof that these puzzles are linked can only be obtained when we finally have a proper consumption-based model to price assets. That will explain why current models failed, something we cannot do today.

References

[1] **Abel, A.** (1990), "Asset prices under habit formation and catching up with the Joneses," *American Economic Review Papers and Proceedings* 80, 38 - 42.

[2] **Bai, J.** (2005), “Panel data models with interactive fixed effects,” Working Paper: New York University.

[3] **Backus, D., Foresi, S. and Telmer, C.** (1995), “Interpreting the forward premium anomaly,” *The Canadian Journal of Economics* 28, 108 - 119.

[4] **Brandt, M. W., Cochrane, J. H. and Santa-Clara, P.** (2006), "International risk sharing is better than you think, or exchange rates are too smooth," *Journal of Monetary Economics* 53, 671 - 698.

[5] **Breeden, D.** (1979), "An intertemporal asset pricing model with stochastic consumption and investment opportunities," *Journal of Financial Economics* 7, 265 - 296.

[6] **Burnside, C. , Eichenbaum, M. Kleshchlski, I. and Rebelo, S.** (2006), "The returns to currency speculation," NBER Working paper #12489.

[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] **Campbell, J. Y. and Shiller, R. J.** (1988), "The dividend-price ratio and the expectations of future dividends and discount factors," *Review of Financial Studies* 1, 195 - 228.

[9] **Chamberlain, Gary, and Rothschild, M.** (1983), "Arbitrage, factor structure, and mean-variance analysis on large asset markets," *Econometrica*, vol. 51(5), pp. 1281-1304.

[10] **Cochrane, J.H.** (2001), "Asset pricing," Princeton Univerity Press.

[11] **Cochrane, J.H.** (2006), "Financial markets and the real economy," Mimeo.

[12] **Cochrane, J.H. and Hansen, L.** (1992), "Asset pricing explorations for macroeconomics," NBER Working paper # 12026.

[13] **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.

[14] **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.

[15] **Connor, G. and Korajczyk, R.** (1993), "A test for the number of factors in am approximate factor structure", *Journal of Finance* 48, 1263 - 1291.

[16] **Constantinides, G. M. and Duffie, D.** (1996), "Asset pricing with heterogeneous consumers," *Journal of Political Economy* 104, 219 - 240.

[17] **Cumby, R. E.** (1988), "Is it risk? Explaining deviations from uncovered interest parity," *Journal of Monetary Economics* 22, 279 - 299.

[18] **Dhrymes, P. J.** (1974), "Econometrics: statistical foundations and applications," New York, Springer-Verlag.

[19] **Engel, C.** (1996), "The forward discount anomaly and the risk premium: a survey of recent evidence," *Journal of Empirical Finance* 3, 123 - 192.

[20] **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.

[21] **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.

[22] **Fama, E. F.** (1984), "Term premiums in bond returns," *Journal of Financial Economics*, 13, 529 - 546.

[23] **Fama, E. F. and Farber, A.** (1979), "Money, bonds and foreign exchange," *American Economic Review* 69, 269 - 282.

[24] **Fama, E. F. and French, K, R.** (1988), "Dividend yields and expected stock returns," *Journal of Financial Economics* 22, 3 - 27.

[25] **Fama, E. F. and French, K, R.** (1989), "Business conditions and the expected stock returns," *Journal of Financial Economics* 25, 23 - 49.

[26] **Fama, E. F. and French, K, R.** (1993), "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics* 33, 3 - 56.

[27] **Fama, E. F. and French, K, R.** (1996), "Multifactors explanations of asset-pricing anomalies," *Journal of Financial Economics* 51, 55 - 84.

[28] **Frankel, J. A.** (1979), "The diversifiability of exchange risk," *Journal of International Economics* 9, 379-393.

[29] **Froot, K. A.** (1990), "Short rates and expected asset returns," NBER Working paper # 3247.

[30] **Gomes, F.A. R. and Issler, J.V.** (2006), "A General test for rule-of-thumb behavior." Paper presented at the European Econometric Society Meeting of Vienna, downloadable form http://www.eea-esem.com/EEA-ESEM/2006/prog/viewpaper.asp?pid=1764

[31] **Hansen, L. P.** (1982), "Large sample properties of generalized method of moments estimators," *Econometrica* 50, 1029 - 1054.

[32] **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.

[33] **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.

[34] **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.

[35] **Hansen, L. and Singleton, K.** (1982), "Generalized instrumental variables estimation of nonlinear expectations models," *Econometrica* 50(5), pp. 1269-1286.

[36] **Hansen, L. and Singleton, K.** (1983), "Stochastic consumption, risk aversion and the temporal behavior of asset returns". *Journal of Political Economy* 91(2), 249-265.

[37] **Hansen, L. and Singleton, K.** (1984), “Erratum of the article: “Generalized instrumental variables estimation of nonlinear expectations models,” *Econometrica* 52(1), pp. 267-268.

[38] **Harrison, J. M. and Kreps, D. M.** (1979), “Martingales and arbitrage in multiperiod securities markets,” *Journal of Economic Theory* 20, 249 - 268.

[39] **Hodrick, R. J.** (1981), “International asset pricing with time-varying risk premiums,” *Journal of International Economics* 11, 573 - 587.

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

[41] **Hodrick, R. J.** (1989), “US international capital flows: perspectives from rational maximizing models,” Carnegie-Rochester Conference Series on Public Policy 30, 1 - 29.

[42] **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.

[43] **Huang, R. D.** (1989), “An analysis of intertemporal pricing for forward exchange contracts,” *Journal of Finance* 44, 183 - 194.

[44] **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.

[45] **Lehmann, B. and Modest, D.** (1988), “The empirical foundations of the arbitrage pricing theory,” *Journal of Financial Economics*, 21, 213-254.

[46] **Lewis, K. K.** (1990), “The behavior of Eurocurrency returns across different holding periods and regimes,” *Journal of Finance* 45, 1211 - 1236.

[47] **Lucas, R. E.** (1978), “Asset pricing in an exchange economy,” *Econometrica*, vol. 46, pp. 1429 - 1445.

[48] **Lucas, R. E.** (1982), “Interest rates and currency prices in a two-country world,” *Journal of Monetary Economics* 10, 335 - 360.

[49] **Lustig, H. and Verdelhan, A.** (2006 a), "The cross-section of foreign currency risk premia and consumption growth risk," *American Economic Review*, forthcoming.

[50] **Lustig, H. and Verdelhan, A.** (2006 b) "Investing in foreign currency is like betting on your intertemporal marginal rate of substitution", *Journal of the European Economic Association*, 4:644-655.

[51] **Mark, N.** (1985), "On time varying risk premiums in the foreign exchange market: an econometric analysis," *Journal of Monetary Economics* 16, 3 - 18.

[52] **Mehra, R. and Prescott, E.** (1985), "The equity premium: a puzzle," *Journal of Monetary Economics* 15, 627 - 636.

[53] **Modjtahedi, B.** (1991), "Multiple maturities and time-varying risk premia in forward exchange markets: An econometric analysis," *Journal of International Economics* 30, 69-86.

[54] **Ross, S. A.** (1976), "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 13, pp. 341-360.

[55] **Shiller R. J.** (1982), "Consumption, asset markets and economic fluctuations," *Carnegie-Rochester Conference on Public Policy*, 17, pp. 203-238.

Appendix

A. Return-Based Estimates of the SDF Mimicking Portfolio

A.1. The Approach of Hansen and Jagganathan (1991)

Given a set of N traded returns stacked in a vector $\mathbf{R}_{t+1}$, Hansen and Jagganathan construct algebraically the unique projection of any SDF on the space of returns. It is given by the following linear combination of traded returns:

$$\overline{M^*_{t+1}} \equiv \mathbf{1}' \left[\mathbb{E}(\mathbf{R}_{t+1}\mathbf{R}'_{t+1})\right]^{-1} \mathbf{R}_{t+1}, \tag{32}$$

where $\mathbf{1}$ is $N \times 1$ a vector of ones. It is straightforward to verify that the estimator of M^*_{t+1} will obey a vector version of the Pricing Equation, since:

$$\mathbb{E}\left\{\overline{M^*_{t+1}}\mathbf{R}'_{t+1}\right\} = \left\{\mathbf{1}'\left[\mathbb{E}(\mathbf{R}_{t+1}\mathbf{R}'_{t+1})\right]^{-1}\mathbb{E}\left[\mathbf{R}_{t+1}\mathbf{R}'_{t+1}\right]\right\} = \mathbf{1}'.$$

Even is all returns are non-negative, it is possible that the projection of the SDF on the space of returns to be negative for some t, although this is not very common empirically.

A.2. The Multifactor-Model Approach

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 variance-covariance matrix of returns, the second one explains the next largest portion, and so on.

Factor models summarize the systematic variation of the N elements of the vector $\mathbf{R}_t = (R_{1,t}, R_{2,t}, \ ..., R_{N,t})'$ using a reduced number of K factors, $K < N$. Consider a K-factor model in R^i_t:

$$R_{i,t} = a_i + \sum_{k=1}^{K} \beta_{i,k} f_{k,t} + \eta_{it} \tag{33}$$

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

$$\text{plim} \frac{1}{N} \sum_{i=1}^{N} \eta_{i,t} = 0.$$

Denote by $\Sigma_{\mathbf{r}} = \mathbb{E}\left(\mathbf{R}_t\mathbf{R}'_t\right) - \mathbb{E}\left(\mathbf{R}_t\right)\mathbb{E}\left(\mathbf{R}'_t\right)$ the variance-covariance matrix of returns. The first principal component of the elements of $\mathbf{R}_t$ is a linear combination $\theta'\mathbf{R}_t$ with maximal variance subject to the normalization that θ has unit norm, i.e., $\theta'\theta = 1$. 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 $\mathbf{R}_t$ are consistent estimates of the $f_{k,t}$'s. Factor loadings can be estimated consistently by simple OLS regressions of the form (33).

To understand why we need the unit-norm condition, recall that the variance of the first principal component $\theta'\mathbf{R}_t$ is $\theta'\Sigma_{\mathbf{r}}\theta$. As discussed in Dhrymes (1974), because we want to maximize this variance, the problem has no unique solution – we can make the variance as large as we want by multiplying θ by a constant $\kappa > 1$. Indeed, we are facing a *scale* problem, which is solved by imposing unit norm, i.e., $\theta'\theta = 1$.

When a multi-factor approach is used, the first step is to specify the number of 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 unique SDF mimicking portfolio M_t^*. For that, we shall rely on a purely statistical model, combined with the Pricing Equation.

Given estimates of the factor model of R_t^i in (33), one can write the respective expected-beta return expression

$$\mathbb{E}(R_i) = \gamma + \sum_{k=1}^{K} \beta_{i,k}\lambda_k, \qquad i = 1, 2, ..., N$$

where λ_k is interpreted as the price of the k-th risk factor. 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 λ coefficients directly by

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

and consequently to estimate only γ via a cross-sectional regression[26].

Given these coefficients, one can easily get an estimate of M_t^*,

$$\widetilde{M_t^*} \equiv a + \sum_{k=1}^{K} b_k f_{k,t}$$

where (a, b) is related to (λ, γ) through

$$a \equiv \frac{1}{\gamma} \text{ and } b \equiv -\gamma \left[cov(ff')\right]^{-1} \lambda,$$

It is easy then to see the equivalence between the beta pricing model and the linear model for the SDF. More, it is immediate that

$$\mathbb{E}(\widetilde{M_t^*} R_{i,t}) = 1, \quad i = 1, 2, ..., N.$$

It is important to stress that we need to impose a scale whenever a factor-model is used. Here, it is implicitly imposed that $VAR\left(\widetilde{M_t^*}\right) = 1$; see Cochrane (2001).

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.[27] We followed Lehmann and Modest (1988) and Connor and Korajczyk (1988), taking the pragmatic view whereby increasing K until the estimate of M_t^* changed very little due to the last increment in the number of factors. We also performed a robustness analysis for the results of all of our statistical tests using different estimates of M_t^* associated with different K's. Results changed very little around our choice of K.

B. Tables and Figures

Table 1: Data summary statistics

International quarterly data: observations from 1990:I to 2004:III

Government bonds	US$ real net return on the covered trading of foreign government bonds		US$ real net return on the uncovered trading of foreign government bonds		US$ real excess return on the uncovered over the covered trading of foreign government bonds	
	Sample Mean (% per year)	Sample S.D. (% per year)	Sample Mean (% per year)	Sample S.D. (% per year)	Sample Mean (% per year)	Sample S.D. (% per year)
British	1,334	1,681	4,911	9,857	3,541	9,421
Canadian	1,746	1,122	1,956	5,517	0,207	5,605
German	0,913	2,824	3,123	11,138	2,195	11,818
Japanese	1,777	1,342	1,691	12,501	-0,084	12,590
Swiss	1,750	1,450	3,049	12,196	1,282	12,344

US quarterly data: observations from 1990:I to 2004:III

US$ real net return on 90-day Treasury-Bill		US$ real net return on S&P500		US$ real excess return on S&P500 over 90-day Treasury-Bill	
Sample Mean (% per year)	Sample S.D. (% per year)	Sample Mean (% per year)	Sample S.D. (% per year)	Sample Mean (% per year)	Sample S.D. (% per year)
1,422	1,098	8,784	16,815	7,285	16,409

Table 2: Testing overidentifying restrictions of consumption models

Testing consumption models taking into account only excess return of S&P500 over 90 day Treasury-Bill

CRRA Preference

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(R_{t+1}^{SP}-R_{t+1}^{b}\right)\right]=0$$

α (p-value)	J-statistic (p-value)
160,252	0,157
(0,003)	(0,181)

Epstein and Zin (1991) Preference

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)}\left(R_{t+1}^{SP}-R_{t+1}^{b}\right)\right]=0$$

α (p-value)	ρ (p-value)	J-statistic (p-value)
181,205	84,562	0,168
(0,015)	(0,248)	(0,086)

Abel (1990) Preference

$$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(\frac{C_t}{C_{t-1}}\right)^{-\kappa(\alpha-1)}\left(R_{t+1}^{SP}-R_{t+1}^{b}\right)\right]=0$$

α (p-value)	κ (p-value)	J-statistic (p-value)
164,622	-0.198	0,147
(0,013)	(0,699)	(0,091)

Testing consumption models taking into account both return on S&P500 and excess return of S&P500 over 90 day Treasury-Bill

CRRA Preference

$$\begin{cases} E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(R_{t+1}^{SP}-R_{t+1}^{b}\right)\right]=0 \\ E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha} R_{t+1}^{SP}\right]=1 \end{cases}$$

β (p-value)	α (p-value)	J-statistic (p-value)
1,009	7,347	0,467
(0.000)	(0.000)	(0.007)

Epstein and Zin (1991) Preference

$$\begin{cases} E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)}\left(R_{t+1}^{SP}-R_{t+1}^{b}\right)\right]=0 \\ E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)} R_{t+1}^{SP}\right]=1 \end{cases}$$

β (p-value)	α (p-value)	ρ (p-value)	J-statistic (p-value)
1,011	3,332	577,616	0,469
(0.000)	(0.000)	(0,855)	(0.004)

Abel (1990) Preference

$$\begin{cases} E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(\frac{C_t}{C_{t-1}}\right)^{-\kappa(\alpha-1)}\left(R_{t+1}^{SP}-R_{t+1}^{b}\right)\right]=0 \\ E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(\frac{C_t}{C_{t-1}}\right)^{-\kappa(\alpha-1)} R_{t+1}^{SP}\right]=1 \end{cases}$$

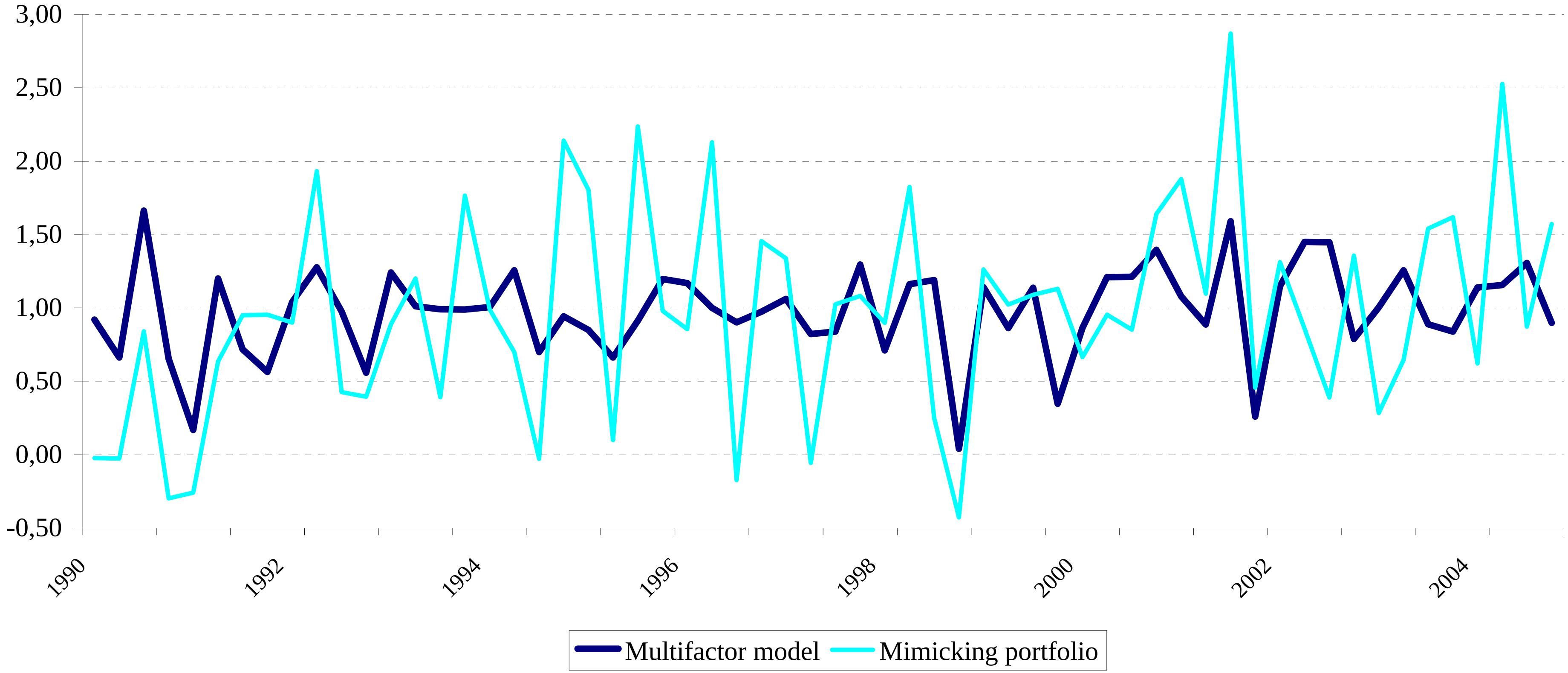
β (p-value)	α (p-value)	κ (p-value)	J-statistic (p-value)
1,009	9,681	0,289	0,418
(0.000)	(0.000)	(0.043)	(0.012)

Notes: Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1990:1 to 2004:3.

Instrument set: real consumption and GDP instantaneous growth rates, consumption/GDP ratio, lagged value of the returns in question and dividend/price ratio and a constant.

Parameters: β is is the one-period discount rate, α is the relative risk aversion coefficient, κ is the time-separability parameter and ρ equals the inverse elasticity of intertemporal substitution.

Figure 1: Pricing kernels with US domestic financial market (1990:I - 2004:III): $\tilde{M}^*_{t+1}$ and $\bar{M}^*_{t+1}$



Pricing kernel estimators summary statistics (quarterly)

	Mean	S.D.	Max.	Min.
Multifactor model	0,977	0,318	1,664	0,041
Mimicking portfolio	0,962	0,725	2,871	-0,428

Table 3: Equity-Premium Puzzle tests (single-equation)

Testing the capacity of our return-based pricing kernels to price excess return of S&P500 over 90-day Treasury-Bill

Instrument sets:	$\frac{C_t}{C_{t-1}};\frac{C_{t-1}}{C_{t-2}};\frac{C_{t-2}}{C_{t-3}};const$		$\frac{Y_t}{Y_{t-1}};\frac{Y_{t-1}}{Y_{t-2}};\frac{Y_{t-2}}{Y_{t-3}};const$		$\frac{C_t}{Y_t};\frac{C_{t-1}}{Y_{t-1}};\frac{C_{t-2}}{Y_{t-2}};const$		$R_t^b;R_{t-1}^b;R_{t-2}^b;const.$		$R_t^{SP};R_{t-1}^{SP};R_{t-2}^{SP};const.$		$(D/P)_t$; $(D/P)_{t-1}$; $(D/P)_{t-2}$; *const*	
Euler equations:	$\mu_1\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	J-stat. (p-value)
$E_t\left[\tilde{M}^*_{t+1}\frac{P_t(i^{SP}_{t+1}-i^b_{t+1})}{P_{t+1}}-\mu_1\right]=0$	0,385	0,028	0,385	0,069	-0,065	0,069	0,385	0,047	0,385	0,056	0,385	0,130
	(0,724)	(0,673)	(0,750)	(0,285)	(0,963)	(0,275)	(0,768)	(0,456)	(0,747)	(0,382)	(0,686)	(0,063)
$E_t\left[\bar{M}^*_{t+1}\frac{P_t(i^{SP}_{t+1}-i^b_{t+1})}{P_{t+1}}-\mu_1\right]=0$	0,659	0,089	0,659	0,087	0,424	0,089	0,659	0,070	0,659	0,075	0,659	0,075
	(0,571)	(0,189)	(0,495)	(0,197)	(0,715)	(0,183)	(0,640)	(0,267)	(0,542)	(0,238)	(0,516)	(0,238)

Testing the capacity of our return-based pricing kernels to price return on S&P500

Instrument sets:	$\frac{C_t}{C_{t-1}};\frac{C_{t-1}}{C_{t-2}};\frac{C_{t-2}}{C_{t-3}};const$		$\frac{Y_t}{Y_{t-1}};\frac{Y_{t-1}}{Y_{t-2}};\frac{Y_{t-2}}{Y_{t-3}};const$		$\frac{C_t}{Y_t};\frac{C_{t-1}}{Y_{t-1}};\frac{C_{t-2}}{Y_{t-2}};const$		$R_t^b;R_{t-1}^b;R_{t-2}^b;const.$		$R_t^{SP};R_{t-1}^{SP};R_{t-2}^{SP};const.$		$(D/P)_t$; $(D/P)_{t-1}$; $(D/P)_{t-2}$; *const*	
Euler equations:	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)
$E_t\left[\tilde{M}^*_{t+1}\frac{P_t(1+i^{SP}_{t+1})}{P_{t+1}}-(1-\mu_2)\right]=0$	-2,014	0,071	-2,014	0,043	-1,272	0,006	-2,014	0,044	-2,014	0,111	-2,014	0,144
	(0,403)	(0,269)	(0,489)	(0,495)	(0,680)	(0,986)	(0,490)	(0,480)	(0,470)	(0,098)	(0,470)	(0,044)
$E_t\left[\bar{M}^*_{t+1}\frac{P_t(1+i^{SP}_{t+1})}{P_{t+1}}-(1-\mu_2)\right]=0$	-0,891	0,094	-0,891	0,046	-0,364	0,010	-0,891	0,117	-0,891	0,053	-0,891	0,087
	(0,875)	(0,169)	(0,879)	(0,470)	(0,959)	(0,903)	(0,905)	(0,087)	(0,892)	(0,407)	(0,902)	(0,191)

Notes: Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1990:1 to 2004:3.

Table 4: Forward-Premium Puzzle tests (single-equation)

Testing the capacity of our return-based pricing kernels to price excess return of uncovered over covered trading of foreign government bonds

$$E_t\left[\tilde{M}^*_{t+1}\frac{P_t(1+i^*_{t+1})}{P_{t+1}}\frac{(S_{t+1}-{}_tF_{t+1})}{S_t}-\mu_1\right]=0$$

	British bonds		Canadian bonds		German bonds		Japanese bonds		Swiss bonds	
	$\mu_1\times 100$	J-stat.	$\mu_1\times 100$	J-stat.	$\mu_1\times 100$	J-stat.	$\mu_1\times 100$	J-stat.	$\mu_1\times 100$	J-stat.
Instrument sets:	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
$\frac{C_t}{C_{t-1}};\frac{C_{t-1}}{C_{t-2}};\frac{C_{t-2}}{C_{t-3}};const$	0,694	0,055	-0,019	0,044	0,686	0,080	-0,306	0,081	0,308	0,077
	(0,182)	(0,398)	(0,964)	(0,486)	(0,493)	(0,224)	(0,708)	(0,220)	(0,703)	(0,265)
$\frac{Y_t}{Y_{t-1}};\frac{Y_{t-1}}{Y_{t-2}};\frac{Y_{t-2}}{Y_{t-3}};const$	0,694	0,030	-0,019	0,030	0,686	0,038	-0,306	0,039	0,308	0,030
	(0,120)	(0,649)	(0,962)	(0,649)	(0,428)	(0,552)	(0,715)	(0,540)	(0,617)	(0,675)
$\frac{C_t}{Y_t};\frac{C_{t-1}}{Y_{t-1}};\frac{C_{t-2}}{Y_{t-2}};const$	1,007	0,025	0,057	0,040	0,879	0,075	0,038	0,069	0,568	0,081
	(0,050)	(0,704)	(0,884)	(0,519)	(0,302)	(0,238)	(0,965)	(0,275)	(0,379)	(0,237)
$R^C_t;R^C_{t-1};R^C_{t-2};const$	0,694	0,016	-0,019	0,054	0,686	0,071	-0,306	0,113	0,308	0,100
	(0,140)	(0,821)	(0,950)	(0,398)	(0,472)	(0,259)	(0,721)	(0,094)	(0,677)	(0,167)
$R^U_t;R^U_{t-1};R^U_{t-2};const$	0,694	0,027	-0,019	0,033	0,686	0,093	-0,306	0,092	0,308	0,048
	(0,089)	(0,679)	(0,958)	(0,605)	(0,464)	(0,167)	(0,720)	(0,171)	(0,640)	(0,475)
$\frac{{}_tF_{t+1}}{S_t};\frac{{}_{t-1}F_t}{S_{t-1}};\frac{{}_{t-2}F_{t-1}}{S_{t-2}};const$	1,007	0,029	0,057	0,114	0,879	0,066	0,038	0,131	0,568	0,085
	(0,076)	(0,655)	(0,878)	(0,092)	(0,240)	(0,300)	(0,966)	(0,061)	(0,376)	(0,222)

$$E_t\left[\bar{M}^*_{t+1}\frac{P_t(1+i^*_{t+1})}{P_{t+1}}\frac{(S_{t+1}-{}_tF_{t+1})}{S_t}-\mu_1\right]=0$$

	British bonds		Canadian bonds		German bonds		Japanese bonds		Swiss bonds	
	$\mu_1\times 100$	J-stat.	$\mu_1\times 100$	J-stat.	$\mu_1\times 100$	J-stat.	$\mu_1\times 100$	J-stat.	$\mu_1\times 100$	J-stat.
Instrument sets:	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
$\frac{C_t}{C_{t-1}};\frac{C_{t-1}}{C_{t-2}};\frac{C_{t-2}}{C_{t-3}};const$	0,840	0,048	-0,442	0,017	0,575	0,056	-1,285	0,057	0,212	0,080
	(0,055)	(0,454)	(0,233)	(0,812)	(0,569)	(0,390)	(0,216)	(0,382)	(0,816)	(0,246)
$\frac{Y_t}{Y_{t-1}};\frac{Y_{t-1}}{Y_{t-2}};\frac{Y_{t-2}}{Y_{t-3}};const$	0,840	0,050	-0,442	0,018	0,575	0,061	-1,285	0,047	0,212	0,070
	(0,067)	(0,438)	(0,234)	(0,798)	(0,549)	(0,350)	(0,287)	(0,462)	(0,798)	(0,318)
$\frac{C_t}{Y_t};\frac{C_{t-1}}{Y_{t-1}};\frac{C_{t-2}}{Y_{t-2}};const$	0,990	0,051	-0,396	0,023	0,669	0,099	-1,092	0,097	0,343	0,106
	(0,090)	(0,423)	(0,274)	(0,728)	(0,479)	(0,143)	(0,353)	(0,151)	(0,670)	(0,144)
$R^C_t;R^C_{t-1};R^C_{t-2};const$	0,840	0,023	-0,442	0,017	0,575	0,043	-1,285	0,103	0,212	0,091
	(0,106)	(0,728)	(0,115)	(0,808)	(0,496)	(0,488)	(0,211)	(0,127)	(0,791)	(0,200)
$R^U_t;R^U_{t-1};R^U_{t-2};const$	0,840	0,101	-0,442	0,039	0,575	0,079	-1,285	0,111	0,212	0,078
	(0,067)	(0,135)	(0,196)	(0,532)	(0,537)	(0,222)	(0,271)	(0,098)	(0,805)	(0,248)
$\frac{{}_tF_{t+1}}{S_t};\frac{{}_{t-1}F_t}{S_{t-1}};\frac{{}_{t-2}F_{t-1}}{S_{t-2}};const$	0,990	0,051	-0,396	0,108	0,669	0,079	-1,092	0,121	0,343	0,090
	(0,090)	(0,423)	(0,254)	(0,107)	(0,473)	(0,222)	(0,250)	(0,079)	(0,628)	(0,204)

Notes: Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1990:1 to 2004:3.

Table 5: Forward-Premium Puzzle tests (single-equation)

Testing the capacity of our return-based pricing kernels to price return on uncovered trading of foreign government bonds

$$E_t\left[\widetilde{M}^*_{t+1}\frac{P_t(1+i^*_{t+1})}{P_{t+1}}\frac{S_{t+1}}{S_t}-(1-\mu_2)\right]=0$$

Instrument sets:	British bonds $\mu_2\times100$ (p-value)	British bonds J-stat. (p-value)	Canadian bonds $\mu_2\times100$ (p-value)	Canadian bonds J-stat. (p-value)	German bonds $\mu_2\times100$ (p-value)	German bonds J-stat. (p-value)	Japanese bonds $\mu_2\times100$ (p-value)	Japanese bonds J-stat. (p-value)	Swiss bonds $\mu_2\times100$ (p-value)	Swiss bonds J-stat. (p-value)
$\frac{C_t}{C_{t-1}};\frac{C_{t-1}}{C_{t-2}};\frac{C_{t-2}}{C_{t-3}};const$	1,703	0,062	2,336	0,063	1,926	0,059	2,625	0,057	2,132	0,068
	(0,524)	(0,342)	(0,349)	(0,334)	(0,487)	(0,366)	(0,299)	(0,382)	(0,473)	(0,332)
$\frac{Y_t}{Y_{t-1}};\frac{Y_{t-1}}{Y_{t-2}};\frac{Y_{t-2}}{Y_{t-3}};const$	1,703	0,037	2,336	0,029	1,926	0,032	2,625	0,042	2,132	0,058
	(0,610)	(0,564)	(0,463)	(0,661)	(0,573)	(0,625)	(0,402)	(0,504)	(0,561)	(0,407)
$\frac{C_t}{Y_t};\frac{C_{t-1}}{Y_{t-1}};\frac{C_{t-2}}{Y_{t-2}};const$	0,197	0,017	1,069	0,018	0,525	0,015	1,086	0,013	0,579	0,027
	(0,961)	(0,808)	(0,778)	(0,794)	(0,900)	(0,835)	(0,779)	(0,862)	(0,898)	(0,702)
$R^C_t;R^C_{t-1};R^C_{t-2};const$	1,703	0,062	2,336	0,042	1,926	0,053	2,625	0,062	2,132	0,082
	(0,594)	(0,333)	(0,450)	(0,497)	(0,598)	(0,406)	(0,377)	(0,333)	(0,577)	(0,233)
$R^U_t;R^U_{t-1};R^U_{t-2};const$	1,703	0,105	2,336	0,049	1,926	0,099	2,625	0,101	2,132	0,097
	(0,588)	(0,119)	(0,456)	(0,439)	(0,559)	(0,143)	(0,391)	(0,135)	(0,606)	(0,178)
$\frac{{}_tF_{t+1}}{S_t};\frac{{}_{t-1}F_t}{S_{t-1}};\frac{{}_{t-2}F_{t-1}}{S_{t-2}};const$	0,197	0,080	1,069	0,059	0,525	0,044	1,086	0,016	0,579	0,038
	(0,960)	(0,218)	(0,616)	(0,357)	(0,869)	(0,480)	(0,741)	(0,821)	(0,878)	(0,577)

$$E_t\left[\overline{M}^*_{t+1}\frac{P_t(1+i^*_{t+1})}{P_{t+1}}\frac{S_{t+1}}{S_t}-(1-\mu_2)\right]=0$$

Instrument sets:	British bonds $\mu_2\times100$ (p-value)	British bonds J-stat. (p-value)	Canadian bonds $\mu_2\times100$ (p-value)	Canadian bonds J-stat. (p-value)	German bonds $\mu_2\times100$ (p-value)	German bonds J-stat. (p-value)	Japanese bonds $\mu_2\times100$ (p-value)	Japanese bonds J-stat. (p-value)	Swiss bonds $\mu_2\times100$ (p-value)	Swiss bonds J-stat. (p-value)
$\frac{C_t}{C_{t-1}};\frac{C_{t-1}}{C_{t-2}};\frac{C_{t-2}}{C_{t-3}};const$	-1,085	0,089	0,123	0,091	-0,807	0,089	1,002	0,088	-0,202	0,061
	(0,853)	(0,189)	(0,983)	(0,181)	(0,893)	(0,189)	(0,854)	(0,193)	(0,976)	(0,385)
$\frac{Y_t}{Y_{t-1}};\frac{Y_{t-1}}{Y_{t-2}};\frac{Y_{t-2}}{Y_{t-3}};const$	-1,085	0,039	0,123	0,045	-0,807	0,045	1,002	0,050	-0,202	0,038
	(0,862)	(0,540)	(0,984)	(0,478)	(0,900)	(0,478)	(0,868)	(0,438)	(0,980)	(0,585)
$\frac{C_t}{Y_t};\frac{C_{t-1}}{Y_{t-1}};\frac{C_{t-2}}{Y_{t-2}};const$	-0,943	0,011	0,370	0,013	-0,615	0,007	1,099	0,005	-0,028	0,007
	(0,901)	(0,890)	(0,960)	(0,862)	(0,938)	(0,940)	(0,883)	(0,962)	(0,997)	(0,946)
$R^C_t;R^C_{t-1};R^C_{t-2};const$	-1,085	0,060	0,123	0,113	-0,807	0,029	1,002	0,071	-0,202	0,146
	(0,883)	(0,349)	(0,987)	(0,094)	(0,867)	(0,655)	(0,891)	(0,259)	(0,979)	(0,052)
$R^U_t;R^U_{t-1};R^U_{t-2};const$	-1,085	0,082	0,123	0,031	-0,807	0,104	1,002	0,083	-0,202	0,046
	(0,887)	(0,210)	(0,986)	(0,630)	(0,915)	(0,123)	(0,882)	(0,206)	(0,981)	(0,490)
$\frac{{}_tF_{t+1}}{S_t};\frac{{}_{t-1}F_t}{S_{t-1}};\frac{{}_{t-2}F_{t-1}}{S_{t-2}};const$	-0,943	0,048	0,370	0,071	-0,615	0,056	1,099	0,003	-0,028	0,071
	(0,911)	(0,447)	(0,962)	(0,259)	(0,924)	(0,382)	(0,881)	(0,982)	(0,997)	(0,300)

Notes: Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1990:1 to 2004:3.

Table 6: Equity-Premium Puzzle tests (system)

Testing the capacity of $\tilde{M}^*_{t+1}$ and $\bar{M}^*_{t+1}$ to price jointly: return on S&P500 and also excess return of S&P500 over risk-free short-term bonds

	$E_t\left[\tilde{M}^*_{t+1}\frac{P_t(i^{SP}_{t+1}-i^b_{t+1})}{P_{t+1}}-\mu_1\right]=0$ and $E_t\left[\tilde{M}^*_{t+1}\frac{P_t(1+i^{SP}_{t+1})}{P_{t+1}}-(1-\mu_2)\right]=0$			$E_t\left[\bar{M}^*_{t+1}\frac{P_t(i^{SP}_{t+1}-i^b_{t+1})}{P_{t+1}}-\mu_1\right]=0$ and $E_t\left[\bar{M}^*_{t+1}\frac{P_t(1+i^{SP}_{t+1})}{P_{t+1}}-(1-\mu_2)\right]=0$		
Instrument set:	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)
$\frac{C_t}{C_{t-1}}$; $\frac{Y_t}{Y_{t-1}}$; $\frac{C_t}{Y_t}$; R_t^{SP}; R_t^b; $\left(\frac{D}{P}\right)_t$; $const$.	-0,016	1,775	0,241	-0,414	1,413	0,245
	(0,985)	(0,438)	(0,309)	(0,587)	(0,796)	(0,292)

Notes: Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1990:1 to 2004:3.

Table 7: Fama-French portfolios pricing tests (system)

Testing the capacity of $\tilde{M}^*_{t+1}$ and $\bar{M}^*_{t+1}$ to price jointly the Fama-French zero-cost portfolios

	$E_t\left[\tilde{M}^*_{t+1}(R_m-R_f)_{t+1}-\mu_1\right]=0$, $E_t\left[\tilde{M}^*_{t+1}HML_{t+1}-\mu_2\right]=0$ and $E_t\left[\tilde{M}^*_{t+1}SMB_{t+1}-\mu_3\right]=0$				$E_t\left[\bar{M}^*_{t+1}(R_m-R_f)_{t+1}-\mu_1\right]=0$, $E_t\left[\bar{M}^*_{t+1}HML_{t+1}-\mu_2\right]=0$ and $E_t\left[\bar{M}^*_{t+1}SMB_{t+1}-\mu_3\right]=0$			
Instrument set:	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	$\mu_3\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	$\mu_3\times100$ (p-value)	J-stat. (p-value)
$\frac{C_t}{C_{t-1}}$; $\frac{Y_t}{Y_{t-1}}$; $\frac{C_t}{Y_t}$; $(R_m-R_f)_t$; HML_t; SMB_t; $const$	-1.167	1.900	-0.054	0,293	-0.702	2.166	0.199	0,253
	(0,165)	(0,003)	(0,925)	(0,707)	(0,326)	(0.000)	(0,684)	(0,829)

Notes: Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1990:1 to 2004:3.

Fama-French portfolios: $(R_m - R_f)$ is the excess return on the market, HML is the average return on two value portfolios minus the average return on two growth portfolios and SMB is the average return on three small portfolios minus the average return on three big portfolios.

Table 8: Forward-Premium Puzzle tests (system)

Testing the capacity of $\widetilde{M}^{*}_{t+1}$ to price jointly: return on uncovered trading of foreign government bonds and also excess return of uncovered over covered trading of foreign government bonds

$$E_t\left[\widetilde{M}^{*}_{t+1}\frac{P_t(1+i^{*}_{t+1})}{P_{t+1}}\frac{(S_{t+1}-{}_tF_{t+1})}{S_t}-\mu_1\right]=0 \quad \text{and} \quad E_t\left[\widetilde{M}^{*}_{t+1}\frac{P_t(1+i^{*}_{t+1})}{P_{t+1}}\frac{S_{t+1}}{S_t}-(1-\mu_2)\right]=0$$

Instrument sets	British bonds			Canadian bonds			German bonds			Japanese bonds			Swiss bonds		
	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)
$\frac{C_t}{C_{t-1}}$; $\frac{Y_t}{Y_{t-1}}$; $\frac{C_t}{Y_t}$; R_t^C; R_t^U; $\frac{{}_tF_{t+1}}{S_t}$; $const.$	1,101 (0,014)	0,657 (0,831)	0,184 (0,555)	0,075 (0,787)	1,606 (0,522)	0,242 (0,305)	0,882 (0,151)	1,072 (0,728)	0,281 (0,189)	0,038 (0,951)	1,642 (0,565)	0,259 (0,241)	0.649 (0,315)	1,093 (0,734)	0,241 (0,378)

Testing the capacity of $\overline{M}^{*}_{t+1}$ to price jointly: return on uncovered trading of foreign government bonds and also excess return of uncovered over covered trading of foreign government bonds

$$E_t\left[\overline{M}^{*}_{t+1}\frac{P_t(1+i^{*}_{t+1})}{P_{t+1}}\frac{(S_{t+1}-{}_tF_{t+1})}{S_t}-\mu_1\right]=0 \quad \text{and} \quad E_t\left[\overline{M}^{*}_{t+1}\frac{P_t(1+i^{*}_{t+1})}{P_{t+1}}\frac{S_{t+1}}{S_t}-(1-\mu_2)\right]=0$$

Instrument sets	British bonds			Canadian bonds			German bonds			Japanese bonds			Swiss bonds		
	$\mu_1\times100$ (p-value)	$\mu_2\times100$ (p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	(p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	(p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	(p-value)	J-stat. (p-value)	$\mu_1\times100$ (p-value)	(p-value)	J-stat. (p-value)
$\frac{C_t}{C_{t-1}}$; $\frac{Y_t}{Y_{t-1}}$; $\frac{C_t}{Y_t}$; R_t^C; R_t^U; $\frac{{}_tF_{t+1}}{S_t}$; $const.$	0,969 (0,016)	0,847 (0,878)	0,202 (0,470)	-0,390 (0,146)	2,134 (0,693)	0,187 (0,540)	0,657 (0,334)	1,166 (0,824)	0,380 (0,039)	1,073 (0,176)	2,849 (0,656)	0,247 (0,284)	0,333 (0,425)	1,877 (0,695)	0,232 (0,412)

Notes: Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1990:1 to 2004:3.

Notes

[1]First draft July 2005. We thank Caio Almeida, Marco Bonomo, Luis Braido, Ricardo Brito, Marcelo Fernandes, Jaime Filho, Luis Renato Lima, Celina Ozawa, Rafael Santos, participants at the XXVII Meeting of the Brazilian Society of Econometrics, the "Econometrics in Rio" meeting, the 2006 Meeting of the Brazilian Finance Society, the 61^{th} European Meeting of the Econometric Society and the seminars at Getulio Vargas Foundation and CAEN for their comments and suggestions. The usual disclaimer applies.

[2]While pricing failures for the CCAPM were found before, it was only with the work of Mehra and Prescott (1985), published the same year as Mark's work, that it became clear that there was something fundamentally wrong with the CCAPM in its canonical form.

[3]We do not claim that returns on equity are not predictable. In fact we know that dividend-price ratios and other variables "predict" returns. The point is that this empirical regularity was not seen, in the early days of research with the CCAPM, as a defining feature of the EPP. Nowadays, however, an empirically successful model ought to take care of this (and many other) non-trivial aspects of asset behavior.

[4]In what follows, we use the terms pricing kernel and stochastic discount factor interchangeably.

[5]If we use all foreign and domestic assets to construct SDF estimates, we should expect to price correctly both the equity and forward premium. If that was the case, we could rule out that explanations based on market imperfections are needed to explain these puzzles, since the existence of an SDF is only guaranteed if the law of one price holds. If we cannot price the equity and forward premium in the exercise, all we can conclude is that the Asset Pricing Equation is inappropriate.

[6]Therefore, the model would only have problems in prcing foreign assets if the residual of such projection was correlated with the part of foreign assets return which has zero price. We do believe that a successful consumption model for domestic markets is very unlikely to present such pattern, but our tests cannot rule it out.

[7]See the comprehensive surveys by Hodrick (1987) and Engel (1996), and the references therein.

[8] In what follows, capital letters are used to represent variables in levels and small letters to represent the logs of these same variables.

[9] The closed-form solution for $\lambda(\cdot)$ is:

$$\lambda(h) = \begin{cases} \frac{1}{h} \times \ln\left[\frac{2\times\left(e^{h}-1-h\right)}{h^{2}}\right], & h \neq 0 \\ 1/3, & h = 0, \end{cases}$$

where $\lambda(\cdot)$ maps from the real line into $(0,1)$.

[10] This is *not* an innocuous assumption. By assuming no arbitrage (stronger than law of one price) we guarantee the existence of a positive M. Uniqueness of M, however, requires complete markets: a very strong assumption. Without uniqueness not all pricing kernels need to be positive.

[11] Of course, one can get directly to (1) when $\alpha_1 = 1$ and $\alpha_0 = 0$ using (2) under log-Normality and Homoskedasticity of $M_t R_{i,t}$. One can also do it from (2) if $[\mathbb{E}_t(z_{U,t+1}) - \mathbb{E}_t(z_{C,t+1})]$ is constant. However, the conditions are very stringent in both cases: there is overwhelming evidence that returns are not log-Normal and homoskedastic, and to think that $[\mathbb{E}_t(z_{U,t+1}) - \mathbb{E}_t(z_{C,t+1})]$ is constant can only be justified as an algebraic simplification for expositional purposes.

Even under log-Normality, if returns are heteroskedastic, $[\mathbb{E}_t(z_{U,t+1}) - \mathbb{E}_t(z_{C,t+1})]$ will be replaced by the difference in conditional variances. Again, this is projection on lagged values of observables, and the same problems alluded above are present.

[12] Frankel (1979) argues that most exchange rate risks are diversifiable, there being no grounds for agents to be rewarded for holding foreign assets.

[13] Here, we are implicitly assuming the absence of short-sale constraints or other frictions in the economy. Our assumption is in contrast with that of Burnside et al. (2006) for whom bid-ask spreads' impact on the profitability of currency speculation plays the main role in generating the FPP.

[14] These asset choices follow Hansen and Singleton (1982, 1984) and Mehra and Prescott (1985).

[15] This is for example the case of Lustig and Verdelhan (2006 b), as they point out in their footnote 8.

[16] Though important in themselves, market imperfections are sometimes invoked to explain the FPP; see Burnside et al. (2006).

[17]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. However, contrary to what we do here, this line of research does not try to relate the EPP and the FPP.

[18]Jumping to our results, we should emphasize that we do not reject (19) for any of the instruments, as well, which means that our SDF satisfies both conditions presented by Backus et al. (1995).

[19]An alternative to the GMM testing procedure described above is to test directly the zero-mean restrictions, instead of estimating by GMM.

[20]See Fama and French (1993) for a complete description of them.

[21]Note that we are not assuming the existence of a risk free rate.

[22]It is straightforward to understand why we obtain these results for CRRA utility. In this case, $M_{t+1} = \left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}$, and GMM chooses α as to make a quadratic form using $M_{t+1}\frac{(i^{SP}_{t+1}-i^{b}_{t+1})P_t}{P_{t+1}} \times z_t$ as close to zero as possible. A large value for α makes $\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha} = \left(\frac{C_t}{C_{t+1}}\right)^{\alpha}$ close to zero, since $\frac{C_t}{C_{t+1}}$ is usually smaller than unity.

[23]Corroborated by the evidence also shown in Table 2.

[24]A more optimistic viewpoint is offered by Cochrane (2006), who contrast the disheartening results of the first years of this research with the recent success stories.

[25]There is obviously a problem with current consumption models for the pricing kernel. Therefore, the profession can only claim to have solved the puzzles when we are finally able to write down a proper working model for aggregate consumption.

[26]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 γ, one set it equal to the real return of the risk free rate.

[27]Despite this relevance, the pure number of factors is not a meaninful question.

CHAPTER 2

On the relative performance of consumption models in foreign and domestic markets.[28]

Abstract

In this paper we follow da Costa and Matos (2007), revisiting the relationship between the Equity and the Forward Premium Puzzles. They are able to show how a same return-based pricing kernel containing sources of risk in the American economy can account for both domestic and international stylized facts. Here, in order to extend this research line, we evaluate the performance of different consumption models in pricing excess returns on the relevant assets, estimating and testing the overidentifying restrictions of Euler equations associated with the Consumption Capital Asset Pricing Model, in its canonical and main reference level versions. Our main finding is that the same (however often unreasonable) values for the parameters are estimated for all models in both equity and foreign exchange markets. In all cases, the rejections or otherwise of overidentifying restrictions occurs for the two markets, suggesting that success and failure stories for the one market repeat themselves in the other one. Our results indicate a strong similarity between the behavior of excess returns in these two markets when modeled as risk premia, providing empirical grounds to believe that the proposed preference-based solutions to puzzles in domestic financial markets can certainly shed light on puzzles in foreign exchange market.

JEL Code: G12; G15. Keywords: Equity Premium, Foreign Currency Risk Premium, CCAPM, Habit Formation, External Time-varying 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 is how Mehra and Prescott (1985) labelled the systematic 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, for reasonable preference parameters.

The FPP, on the other hand, is related to the widely reported departure from the intuitive proposition that the expected return to speculation in the forward foreign exchange market should be zero, conditional on available information.[29]

More important for our purposes is the research line that investigates this puzzle within the framework of the consumption capital asset pricing model – CCAPM. Using a non-linear GMM approach, Mark (1985) and Hodrick (1989), evidence the failure of the canonical version of consumption models to generate a pattern of prices for risk that could explain the FPP being only related to the implicit risk neutrality assumption.

These similar results for both puzzles were never properly linked and despite the skepticism about risk based explanations for the FPP from part of the literature, da Costa and Matos (2007) revisit the relationship between these puzzles, asking whether they deserve (or not) two distinct agendas. Their main finding is that a same pricing kernel is able to account for the equity and the forward premium, which can be viewed as new evidence supporting the usefulness of reuniting the research agendas on these puzzles, since they are two symptoms of the same illness: the incapacity of existing consumption-based models to generate the implied behavior of a pricing kernel that correctly prices asset returns.

In these twenty years of research, one can observe for one side, the CCAPM's inability, in its canonical version, to account for the EPP and the FPP, but for the other side, the evidence that return-based pricing kernels, which are an upper bound on any consumption model – in the sense the pricing capability –, even containing sources or risk in U.S. assets alone, are able to show no signs of the EPP nor the FPP.

In this context, this paper, takes the discussion of this previous paper one step further by evaluating the performance of different models in pricing excess returns on the relevant asstes for each market. Our main 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.

Two are the main features of our procedure, which is the key innovation of this paper. First, we select a representative class of consumption-based asset pricing models pointed as the most promising route to better understand the EPP. This class is comprised by the canonical CCAPM and by other models which look at extending this existing parametric utility function by allowing for relative consumption [Abel's (1990) *catching up with the*

Joneses model]; generalized expected utility [Epstein and Zin's (1991) recursive utility specification]; and the agent to derive utility from the level of consumption relative to this benchmark in ratio and in difference in a Campbell and Cochrane's (1999) slow-moving habit formation approach.

We then estimate and test each one of these models using the Generalized Method of Moments – GMM – for the equity and the foreign exchange markets separately, and also for both markets jointly. For the equity market we test the excess return of the S&P500 over the Treasury bill, while for the foreign exchange markets we consider the excess return on the uncovered over the covered trading of British, Canadian, German and Japanese government bonds jointly.

Second, as for the instruments used to generate the over-identifying restrictions that are tested in our procedure, for both markets, we use exactly the same sets: a first one comprised of lags of macroeconomic instrumental variables – proven to be useful in forecasting the SDF – and lags of the real returns in question as the financial variables; and a second set which replaces these lagged returns by variables with recognized predictability power, as the Fama/French stock-market factors, the factor which captures the one-year momentum anomaly reported in Jegadeesh and Titman's (1993) and the current value for the forward premium.

Briefly, according to our results: *i)* as ever, extremely high values for risk aversion are required to account for risk premia in question; *ii)* despite the significant role played by the past aggregate per capita consumption growth, its inclusion as a reference level in the utility function is far from a reasonable solution*; iii)* breaking the tight link between the coefficient of relative risk aversion and the elasticity of intertemporal substitution alone does not seem to explain none of the puzzles, but *iv)* when besides this separation, is also considered a slow-moving external habit, then for both markets we find a persuasive support for the hypothesis that the agent derives utility from the level of consumption relative to this benchmark not only in ratio, but also in difference.

In particular, we show that the modified version of the model developed by Campbell and Cochrane (1999) with acceptably low values for the utility curvature parameter 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.

These results suggest the strong similarity between the behavior of these excess returns when modeled as risk premia, allowing us to evidence the common pattern of the equity and foreign currency risk premia under a preference-based approach. This finding is robust across different overidentifying restriction sets, which includes different instrument sets besides considering the excess returns not only of the market in question but also of both markets jointly, leading us to believe having provided empirical grounds to support that the proposed preference-based solutions to puzzles in domestic financial markets can certainly shed light on the FPP.

The remainder of the paper is organized as follows. Section 2 gives an account of the literature that tries to relate the two puzzles. More specifically, 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, while subsubsection 2.2.2 describes in details the features shared by most external habit formation models that address the EPP. In section 3, we perform the empirical exercise, estimating with GMM the Euler equations derived from the consumption models in question. Section 4 analyzes the results in order to support that the equity and the currency foreign risk premia have a similar behavior under a preference-based approach. Concluding remarks are offered in section 5.

2. Economic Theory: the equity and the foreign currency risk premia literature

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,

$$1 = \mathbb{E}_t \left(M_{t+1} R^i_{t+1} \right), \tag{34}$$

and

$$0 = \mathbb{E}_t \left[M_{t+1} \left(R^i_{t+1} - R^j_{t+1} \right) \right]. \tag{35}$$

Here, $\mathbb{E}_t(\cdot)$ denotes the conditional expectation given the information available at time t, R^i_{t+1} and R^j_{t+1} represent, respectively, the real gross return on assets i and j at time $t+1$ and M_{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 all assets returns.

An immediate consequence of (35) is that excess returns are explained by the way in which returns covary with the SDF, in fact, if we take R^j to be the return on the risk-free asset, R^f, equation (35) can be expanded to read

$$\mathbb{E}_t\left(R_{t+1}^i\right) - R_t^f = \frac{-cov_t\left(R_{t+1}^i, M_{t+1}\right)}{\mathbb{E}_t\left(M_{t+1}\right)}. \tag{36}$$

It is then, usual practice to interpret the excess expected return as a risk premium and the covariance in the right hand side of (36) 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 mean? Without a model that generates M_{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 financial economics exactly to find a model capable of generating M_{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_{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).[30]

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, the first order conditions for the choice of an optimal consumption and portfolio of this investor, with exogenously specified utility function who faces exogenous asset return process, yield

$$1 = \beta \mathbb{E}_t \left[\frac{u'(C_{t+1})}{u'(C_t)} R^i_{t+1} \right] \quad \forall i \tag{37}$$

and, consequently,

$$0 = \mathbb{E}_t \left[\frac{u'(C_{t+1})}{u'(C_t)} \left(R^i_{t+1} - R^j_{t+1} \right) \right] \quad \forall i, j, \tag{38}$$

where $\beta \in (0, 1)$ is the subjective time discount factor in the representative agent's utility function and, in equilibrium, C_t represents aggregate consumption at time t.

Equations (37) and (38) are consistent with equations (34) and (35) if we define the SDF at time $t+1$, as

$$M_{t+1} \equiv \beta \frac{u'(C_{t+1})}{u'(C_t)}. \tag{39}$$

At first, and considering the very restrictive conditions under which aggregation is possible, one may wonder whether such model is not too off the mark to be taken seriously. It turns out that – see Constantinides (1982) – if markets are complete and agents have von-Neumann Morgenstern preferences, aggregation is not needed for the existence of a representative agent.[31] This is true even if individuals are heterogeneous in preferences and in levels of wealth. Moreover, the representative individual's relative risk aversion is no larger than the most risk averse individual and no smaller than the least risk averse individual. The logic is easy to grasp. Perfect risk sharing, which is what complete markets allow for, requires equalization of marginal growth of consumption across agents in all states of the world.

2.2.1. The canonical consumption-based model and the asset pricing puzzles

Once we have decided that we are going to rely on a representative agent framework, a first important issue to deal with is the specification of utility for this investor. The observation that interest rates and risk premiums have remained stationary over more than a century of US economic growth suggests the use of a few specifications among which the most popular is the time-separable power utility defined over aggregate consumption, C_t,[32]

$$u(C_t) = \frac{C_t^{1-\alpha}}{(1-\alpha)}. \tag{40}$$

In this case, equations (37) and (38) can be rewritten as

$$1 = \beta \mathbb{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} R^i_{t+1} \right] \quad \forall i \tag{41}$$

$$0 = \mathbb{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} \left(R^i_{t+1} - R^j_{t+1} \right) \right]. \tag{42}$$

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.[33]

The consumption Euler equations (41) and (42) are useful to access the EPP, if one takes assets i and j to be $S\&P500$ 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

$$R^C_{t+1} = \frac{{}_tF_{t+1}(1+i^*_t)P_t}{S_t P_{t+1}} \qquad \text{and} \qquad R^U_{t+1} = \frac{S_{t+1}(1+i^*_t)P_t}{S_t P_{t+1}}, \tag{43}$$

where ${}_tF_{t+1}$ and S_t are the forward and spot prices of foreign currency in terms of domestic currency, P_t is the dollar price level and i^*_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

$$0 = \mathbb{E}_t \left\{ \left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} \frac{P_t(1+i^*_t)[{}_tF_{t+1} - S_{t+1}]}{S_t P_{t+1}} \right\}, \tag{44}$$

along the lines of Lucas' (1978) model.

Most papers in the literature estimate the Euler equation of the excess return and test its overidentifying restrictions using the Hansen's (1982) Generalized Method of Moments (GMM).[34] With regards to the FPP, the results obtained in Mark (1985) and Hodrick (1989) report values of the coefficient of relative risk aversion, α, above 40 and 60, respectively, while Engel (1996) reports some estimates for α 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 Mehra and Prescott (1985).

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 attention to the fact that in both cases an extremely high coefficient of risk aversion is estimated. Part of our story is to ask whether the same values for the parameters of different models are needed to account for the risk premia in these markets. This is not the whole story, however. Even if we are willing to accept the "unrealistic" values for the relevant parameters, the overidentifying restrictions must be tested. 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 premia 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.

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 it 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?

In da Costa et al. (2006) we have shown that the same M_{t+1} containing sources of risk in the American economy 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.

In accordance with the mets widely reported in the literature, stock markets commonly show familiar patterns. When we observe an annualized Sharpe ratio for the US stock market, in most cases it ranges from 0.40 to 0.50. Shiller (1982), Hansen and Jagannathan (1991) and Cochrane and Hansen (1992) relate the EPP to the volatility of the pricing kernel required to match this Sharpe ratio.

Since for the foreign exchange market the observed Sharp ratios assume in general quite lower values, what should we expect if we were able of writing down a model that consistently generated a SDF that accounts for the EPP? Would this model also explain the FPP?

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_{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 et al. (2006).

Given our purposes, of simply relating the risk premia of these markets, 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 and promising 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_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,[35] for simplicity, since our intent is to explore two other common features of the habit framework when dealing with equity and foreign currency risk premia.

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

$$u(C_t, X_t) = \frac{1}{(1-\alpha)(1-\varphi)} \left(\frac{C_t}{X_t}\right)^{1-\alpha} X_t^{1-\varphi}, \tag{45}$$

where α is the curvature parameter for relative consumption, while φ assumes this same function for the benchmark level.

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

$$u(C_t, X_t) = \frac{1}{(1-\alpha)(1-\varphi)} \left(C_t - X_t\right)^{1-\alpha} X_t^{\alpha-\varphi}. \tag{46}$$

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 The Abel's (1990) catching up with the Joneses 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 $\overline{C}_{t-1}$.

It is easy to see this formulation as a special case of the equation (45), where $\varphi = 1$ and the level habit is given by $X_t = \overline{C}_{t-1}^{\kappa}$, which enables us to rewrite it as

$$u(C_t) = \frac{1}{(1-\alpha)} \left(\frac{C_t}{\overline{C}_{t-1}^{\kappa}} \right)^{1-\alpha} \tag{47}$$

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

$$M_{t+1}^{CJ} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} \left(\frac{C_t}{C_{t-1}} \right)^{\kappa(\alpha-1)} \tag{48}$$

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 κ. 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, i.e., if the agent is unhappy or happy when an average person did well last period, respectively.

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 (40) 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 a too low intertemporal marginal rate of substitution. As a consequence, the risk-free puzzle arises 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 Weil (1989) and Epstein and Zin (1989, 1991) 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; and *ii)* disentangle the coefficient of risk aversion and the elasticity of intertemporal substitution as in Abel (1999). This utility specification also provides an amplified permanent component of the pricing kernel, satisfying the lower bound for the size of this term derived in Alvarez and Jermann (2002). According to them, this is a characteristic criticism of pricing kernels derived as a function of only consumption data.

The Epstein-Zin objective function can be written as

$$u(C_t, \mathbb{E}_t(U_{t+1})) = \left[(1-\beta)C_t^{\frac{1-\alpha}{\vartheta}} + \beta(\mathbb{E}_t(U_{t+1}^{1-\alpha}))^{\frac{1}{\vartheta}}\right]^{\frac{\vartheta}{1-\alpha}}, \tag{49}$$

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.

To the intent to verify whether the data are consistent with the first order conditions derived from this objective function, we deal with the problem of the unobservality of the period $t+1$ level of utility of the representative agent, following the representation derived in Epstein and Zin (1991).

This approach, whose key is the link between the reference level and the return on the market portfolio, is observationally equivalent to the ratio habit formation model given by equation (45) if one assumes that

$$\log\left(\frac{X_{t+1}}{X_t}\right) = \frac{1}{1-\varphi}\log(R_{M,t+1}) + \omega, \tag{50}$$

where ω is a constant and $R_{M,t+1}$ is the US real gross return on the market portfolio.[36]

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

$$M_{t+1}^{EZ} = \beta\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)}. \tag{51}$$

Slow-moving habit formation approach 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 an 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, let $S_t \equiv (C_t - X_t)/C_t$ 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, η_t, is now time-varying and it can be related to the surplus consumption ratio by $\eta_t \equiv \alpha/S_t$, 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 $\overline{S}_t \equiv (\overline{C}_t - X_t)/\overline{C}_t$. Once again, note that in equilibrium the relations $C_{t+1} \equiv \overline{C}_{t+1}$ and $S_{t+1} \equiv \overline{S}_{t+1}$ hold, which enables us to rewrite the SDF as

$$M_{t+1}^{CC} = \beta \left(\frac{C_{t+1} S_{t+1}}{C_t S_t} \right)^{-\alpha}. \tag{52}$$

Second, with regards the consumption growth, this process is modeled as an independently and identically distributed (i.i.d.) lognormal process,[37]

$$c_{t+1} - c_t = g + \nu_{t+1}, \quad \nu_{t+1}\tilde{\ } iidN(0, \sigma^2), \tag{53}$$

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_t responds to the history of aggregate consumption $\overline{C}$, let the log surplus consumption ratio $\bar{s}_t = \ln(\overline{S}_t)$ evolves as a heteroskedastic AR(1) process,

$$\bar{s}_t = (1 - \phi)\bar{s} + \phi \bar{s}_t + \lambda(\bar{s}_t)(\bar{c}_{t+1} - \bar{c}_t - g) \tag{54}$$

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.[38] The $\lambda(s_t)$ term, labeled sensitive function, is specified following

$$\lambda(s_t) = \begin{cases} \frac{1}{\overline{S}}\sqrt{1-2(s_t-\bar{s})-1}, & s_t \leq s_{\max} \\ 0, & s_t \geq s_{\max} \end{cases} \tag{55}$$

$$\text{where} \quad s_{\max} \equiv \bar{s} + \frac{1}{2}(1-\overline{S}) \text{ and } \overline{S} = \sigma\sqrt{\frac{\alpha}{1-\phi}}$$

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.

The modified 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 in difference, while the traditional habit formation ones take into account the consumption relative to the benchmark only in ratio.

However, in a recent paper, Garcia et al. (2006) propose an extension of Campbell and Cochrane (1999), which nests the Epstein and Zin (1989) model with a habit formation model in difference. Besides allowing for a time-varying risk-free interest rate and breaking the tight link between the coefficient of relative risk aversion and the elasticity of intertemporal substitution, this extension also allows for the absolute value of the reference level to play an independent role in the utility function.

In order then to deal with this disjoint choice of a ratio or a difference model, we follow Garcia et al. (2006), in the sense of proposing not a generalization but a modification of Campbell and Cochrane (1999) model. Since this more general Garcia et al.'s approach corresponds to the equation (46), one must note that with $\varphi = \alpha$, we get the Campbell and Cochrane model, while if $\varphi = 1$, we obtain exactly a preference specification 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.[39]

In our setting, the SDF can be written as

$$M_{t+1}^{MCC} = \beta \left(\frac{C_{t+1} S_{t+1}}{C_t S_t} \right)^{-\alpha} \left(\frac{X_{t+1}}{X_t} \right)^{(\alpha-1)}. \tag{56}$$

3. Empirical application

In this section, we estimate and test the overidentifying restrictions of the Euler equations derived from the 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 bond, second for the excess returns on the uncovered over the covered trading of Canadian, German, Japanese and British government bonds jointly and finally for both markets, considering all these excess returns jointly.

3.1. Pricing test

In order to evidence the common pattern displayed by the equity and the foreign currency risk premia based on a representative agent approach, we start by testing and estimating the Euler equation derived from the standard consumption model. To better understand the behavior of the equity premium, the pricing error on the excess return on the $S\&P500$ over the short-term American government bond is minimized with the Hansen's (1982) Generalized Method of Moments (GMM). For this case the Euler equation is given by

$$\mathbb{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} (R_{t+1}^{SP} - R_{t+1}^{Tb}) \right] = 0. \tag{57}$$

If the purpose is, however, to analyze the behavior of the foreign currency risk premia, then the pricing errors on the excess return on the uncovered over the covered trading of British, Canadian, German and Japanese government bonds are jointly minimized with the Hansen's (1982) GMM. We, therefore, apply the same estimation procedure to the following Euler equations

$$\mathbb{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} (\mathbf{R}_{t+1}^{U} - \mathbf{R}_{t+1}^{C}) \right] = \mathbf{0}. \tag{58}$$

Finally, working as our benchmark, we also consider the case where the pricing errors on the excess return on the equity over the T-bill and on the uncovered over the covered trading of foreign government bonds are jointly minimized using the same econometric technique. This key innovation allows us to analyze how the "interaction" of these markets affects the pattern of each market when considered individually, besides improving our robustness check.

The typical procedure adopted here – 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 the canonical model are reported in Table 2.

One must note that a similar strategy can be performed for any system of Euler equations, allowing us to implement this econometric testing procedure to examine whether the external reference level consumption-based models can account for the EPP and for the FPP.

As before, in order for *a successful model*, we must then have plausible and significant values for the parameters and the over-identifying restriction $T \times J$ test in Hansen (1982) should not reject them. This

To analyse the *catching up with the Joneses* model, the system of conditional moment restrictions is given by

$$E_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} \left(\frac{C_t}{C_{t-1}} \right)^{\kappa(\alpha-1)} (R_{t+1}^{SP} - R_{t+1}^{Tb}) \right] = 0, \tag{59}$$

and

$$E_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\alpha} \left(\frac{C_t}{C_{t-1}} \right)^{\kappa(\alpha-1)} (\mathbf{R}_{t+1}^{U} - \mathbf{R}_{t+1}^{C}) \right] = \mathbf{0}. \tag{60}$$

The results for these models are reported respectively in Table 3.

Now, consider the system of the Euler equations derived from the Epstein and Zin utility specification, which are given by

$$\mathbb{E}_t \left[\left(\frac{C_{t+1}}{C_t} \right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)} (R_{t+1}^{SP} - R_{t+1}^{Tb}) \right] = 0, \tag{61}$$

and

$$\mathbb{E}_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)}(\mathbf{R}_{t+1}^U - \mathbf{R}_{t+1}^C)\mathbf{0}\right] = \mathbf{0}. \tag{62}$$

For this case, the procedure adopted is the same one already described with an additional specificity, how to obtain a time series for the market portfolio. 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 4.

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

$$\mathbb{E}_t\left[\left(\frac{C_{t+1}S_{t+1}}{C_tS_t}\right)^{-\alpha}\left(\frac{X_{t+1}}{X_t}\right)^{(\alpha-1)}(R_{t+1}^{SP} - R_{t+1}^{Tb})\right] = 0, \tag{63}$$

and

$$\mathbb{E}_t\left[\left(\frac{C_{t+1}S_{t+1}}{C_tS_t}\right)^{-\alpha}\left(\frac{X_{t+1}}{X_t}\right)^{(\alpha-1)}(\mathbf{R}_{t+1}^U - \mathbf{R}_{t+1}^C)\right] = \mathbf{0}. \tag{64}$$

When we use this last specification, our main interest remains to evidence the common pattern displayed by the equity and the foreign currency risk premia. 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_t. In order to compute the time series of this process, we need to set its initial value and the values of the parameters σ, g , ϕ and α.

With regards to the initial value of the surplus consumption ratio, we set it at the steady state value, $s_o = \bar{s}$. In choosing the parameters σ, g and ϕ, we follow Campbell and Cochrane (1999), calibrating them. For the risk aversion coefficient α, we follow Garcia et al. (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 α varies between 0.00 and 50.00. The results for this model are reported in Table 5.

Instruments There seems to be a consensus in the return forecasting literature about the rejection of the time-invariant excess returns hypothesis. However, the question of which variables can be considered as good predictors for returns is still open.

Hence, the choice of a representative set of forecasting instruments plays certainly an important role and highlights the relevance of the conditional tests.

First, taking into account the fact that expected returns and business cycles are correlated, as documented in Fama and French (1989), we use the three lags of the following macroeconomic variables: real consumption instantaneous growth rates and real GDP instantaneous growth rates.

However, since in our exercise the forecasting variables, the excess returns are all based on market prices, the association with these conventional macroeconomic variables seems to be not much interesting or informative. Then, we carefully choose the financial instruments, believing that each one of them is very useful in forecasting the variables in question.

Initially, we follow Mark (1985) using a first instrument set (IS1) comprised of the macroeconomic variables already mentioned besides two lags of the real returns in question.[40] In choosing these instrumental variables, we are considering that, for any financial market, one should not omit the possibility that returns could be predictable from past returns.

Next, observing the extensive return forecasts literature, if there seems to be a consensus about the rejection of the time-invariant excess returns hypothesis, on the other hand, the quest of which variables can be considered as good and adequate forecasters remains under discussion. In this sense, the choice of a representative set of forecaster instruments plays certainly a big role, not dismissing this "return forecastability phenomenum" and consequently raising the relevance of the conditional tests.

As regards the FPP, we use the current value for the forward premium, since the well documented predictability power of this variable defines features of this puzzle. In order to analyze the EPP, we follow the 4-factor Cahart's (1997) model, using the factor which captures the one-year momentum anomaly reported in Jegadeesh and Titman's (1993), besides considering the Fama/French stock-market factors: i) $Rm-Rf$, the excess return on the market, is the value-weighted return on all NYSE, AMEX, and NASDAQ stocks minus the one-month Treasury bill rate; *ii)* SMB (Small Minus Big) is the average return on three small portfolios minus the average return on three big portfolios and *iii)* HML (High Minus Low) is the average return on two value portfolios minus the average return on two growth portfolios.[41] These factors are able to summarize the size and the value effects, explaining average returns on stocks and bonds.

The instrument set (IS2) is then comprised of the lagged value of these factors, the current value of the own forward premia besides the same macroeconomic variables.

4. Results

4.1. Data and Summary Statistics

In principle, whenever econometric or statistical tests are performed, it is preferable to employ a large data set either in the time-series (T) or in the cross-sectional dimension (N). Regarding the FPP, the main limitation is the fact that the Chicago Mercantile Exchange, the pioneer of the financial-futures market, only launched currency futures in 1972. In addition to that, only futures data for a few developed countries are available since then. In order to have a common sample for the largest set of countries possible, we considered here U.S. foreign-exchange data for Canada, Germany, Japan and the U.K., covering the period from 1977:1 and ending in 2004:3, on a quarterly frequency. In order to extend the time span of used here, keeping a common sample for all countries, we would have to accept a drastic reduction in the number of countries, which we regard as an inferior choice.

Spot and forward exchange-rate returns were transformed into U.S.$ real returns using the consumer price index in the U.S. 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. To study the EPP we used the U.S.$ real returns on the S&P500 and on 90-day T-Bill. Real returns were obtained using the consumer price index in the U.S.

All macroeconomic variables used in econometric tests were extracted from FED's FRED database. We also employed additional forecasting financial variables that are specific to each test performed, and are listed in the appropriate tables of results. The Jegadeesh and Titman's momentum factor and the Fama-French benchmark factors series were extracted from the French data library.

Table 1 presents summary statistics of our database over the period 1977:1 to 2004:3. The average real return on the S&P500 is 8.95% at an annual rate, while that of the 90-day T-Bill is 1.77%. Real stock returns are much more volatile than the US Treasury Bill return – annualized standard deviation of 16.66% and 1.61% respectively. Over the same period, the real return on covered trading of foreign bonds show means and standard deviations quite similar to that of the U.S. Treasury Bill, reflecting the covered interest rate parity in a frictionless economy. Regarding the return on uncovered trading, means range from 2.49% to 4.78%, while standard deviations range from 5.52% to 13.32%.

We computed the Sharpe ratio for the U.S. stock market to be 0.44, while the Sharpe ratio of the uncovered trading of foreign bonds ranges from 0.04 to 0.28.

4.2. Euler equations

According to results corresponding to the estimates and tests of the overidentifying restrictions of the Euler equations (57) and (58), reported in Table 2, the canonical consumption model 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; $\hat{\alpha} \in [77.93, 113.99]$.

This well-known empirical failure is clearly due to 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.[42]

Analyzing the more complex preference specifications, one can observe Table 3 which presents the findings associated with Euler equations (59) and (60) derived from the Abel (1990) specification. Comparing both equity and foreign exchange markets: *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; *ii)* we have the same magnitude order of the parameter κ and *iii)* contrary to the ratio models, but in accordance with Garcia et al. (2006), we find support for the hypothesis that the absolute value of the past aggregate consumption enters the utility function.

This model is not supported by the equity market nor by the foreign currency market data, since the representative agent is of a jealous sort. Even with a positive $\hat{\kappa}$ – 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 these puzzles with this model.

These main findings are corroborated by the robustness check, which also includes the estimate of the Abel's (1999) framework that captures the notion that the agent's benchmark level depends on current and recent levels of consumption per capita.[43]

Next the results reported in table 4, corresponding to the Epstein and Zin (1991) framework, which adopts a reference level as a function of the return on the market portfolio. These results 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 $\hat{\rho} < \hat{\alpha}$, which is a necessary condition to fit the patterns of the risk premiua, both puzzles are still more exacerbated. We also obtain close values for the elasticity of intertemporal substitution, but these values are not large enough, ranging from 0.011 to 0.033.

Our evidence about the incapacity of the Epstein and Zin (1989) framework to explain the FPP corroborates Colacito and Croce (2005), while regarding this same incapacity to explain the EPP, parts way from Epstein and Zin (1991) and Garcia (2006).[44]

Up to this point, the models which adopt state variables as a function only of societal levels of consumption, besides the Epstein and Zin framework are unable to support the data with reasonable values for the parameters, an evidence robust across different overidentifying restriction sets .

However, this is not a "problem" for us, since our goal is not to find a model that fully explains all asset pricing puzzles. More important for us is that: "these unsuccessful models are all alike", in the sense that, although not necessarily the empirical failures of these models are due to a same reason, when each one of these models fail, it always occurs in a same way for both markets.

Our last attempt to model the risk premia in these different markets relies on a modification 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 difference forms. Although the original version of this setting 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.

In choosing the parameters σ, g and ϕ, we follow Campbell and Cochrane (1999), calibrating them. Since $\hat{\sigma}$ and $\hat{g}$ 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 ϕ, we take it to match the serial correlation of log price/ dividend ratio ratios, obtaining a value of 0.989.

Next, we adopt a different econometric technique to evidence if the risk premia in the equity and the foreign currency markets have a similar behavior, proposing a GMM testing procedure to see if these versions of this habit formation model are supported by the data, instead of calibrating them.

According to the results related to the Euler equations (63) and (64) reported in table 5, we can say that this modified model – with low interest rates, roughly i.i.d. consumption growth with small volatility and acceptably low values of utility curvature – 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 econometric techniques used and the models tested are different. At any rate, with regards the EPP, our values of $\hat{\alpha}$ range from 3.13 to 8.11, quite larger than 0.31 obtained in Garcia et al. (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 $\hat{\alpha}$ range from 2.89 to 7.32, while his values range from 5.60 to 8.20.

As a robustness check for this specific model, we also perform these tests for the original version of the Campbell and Cochrane model. Although, in some cases $\hat{\alpha}$ does not seem to be significant at the 5% level when the first instrument set is used,[45] the results support that not only the unsuccessful models, but also the successful ones are all alike.

Once more the equity and the foreign currency risk premia display common patterns, an evidence now robust to the "class of slow-moving habit formation models".

In general terms: *i)* according the Hansen's test of over-identifying restrictions, none of the models is rejected statistically and *ii)* for most models, we do not have problems with the significance of any parameter estimated.[46]

It is worth mentioning that in all cases the robustness check always confirms the main findings. First, the different but very relevant role played by the instrument sets, mainly the second one. Next, the additional test performed taking into account the overidentifying restrictions characteristic of both markets, whose results are reported in the third column of each table. According to the results, we can argue that for the equity market, the inclusion of the overidentifying restrictions characteristic of the foreign exchange market, not only does not seem to worsen the performance of any consumption-based model, as, in most cases, it still improves this performance, which becomes clear observing the figures 1 and 2. It is also possible to obtain this same evidence for the foreign exchange market.

In order to elucidate the figures, the *square* represents the capacity of the model in question to account for the excess returns, considering the results reported in the third column of each table, i.e., taking into account the overidentifying restrictions characteristic of both markets jointly, while the *ball* represents this same performance but considering the results reported in the first and second columns, which takes into account only the overidentifying restrictions characteristic of the specific market.

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 premia 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 premia, 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 modified version of the slow-moving external habit formation model proposed by Campbell and Cochrane (1999), 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, is useful to explain both puzzles.

References

[1] **Abel, A.** (1990), "Asset prices under habit formation and catching up with the Joneses," *American Economic Review Papers and Proceedings* 80, 38 - 42.

[2] **Abel, A.** (1999), "Risk premiums and term premiums in general equilibrium," *Journal of Monetary Economics* 43, 3 - 33.

[3] **Alvarez, F. and Jermann, U.** (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.

[5] **Burnside, C. , Eichenbaum, M. Kleshchlski, I. and Rebelo, S.** (2006), "The returns to currency speculation," *NBER Working paper n^o* 12489.

[6] **Cahart, M.** (1997), "On persistence in mutual fund performance," *Journal of Finance* 52(1), 57 - 82.

[7] **Campbell, J. and Cochrane, J.** (1999), "By force of habit: a consumption-based explanation of aggregate market behavior," *Journal of Political Economy* 107, 205 - 251.

[8] **Cochrane, J.** (2001),"Asset Pricing," Princeton University Press.

[9] **Cochrane, J. and Hansen, L.** (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.** (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., Issler, J. and Matos, P.** (2007), "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] **Fama, E. and French, K.** (1993), "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics* 33, 3 - 56.

[18] **Fuhrer, J.** (2000), "Habit formation in consumption and its implications for monetary policy models," *American Economic Review* 90, 367 - 390.

[19] **Garcia, R., Renault, É. and Semenov, A.** (2006), "A consumption CAPM with a reference level," Mimeo.

[20] **Gorman, W.**(1953), "Community preference fields," *Econometrica* 21, 63-80.

[21] **Grossman, S. and Shiller, R.** (1981), "The determinants of the variability of stock market prices," *American Economic Review* 71, 222 - 227.

[22] **Hansen, L.** (1982), "Large sample properties of generalized method of moments estimators," *Econometrica* 50, 1029 - 1054.

[23] **Hansen, L. and Richard, S.** (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 Political Economy* 99, 225 - 262.

[25] **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.

[26] **Harrison, J. and Kreps, D.** (1979), "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory* 20, 249 - 268.

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

[28] **Hodrick, R.** (1989), "US international capital flows: perspectives from rational maximizing models," Carnegie-Rochester Conference Series on Public Policy 30, 1-29.

[29] **Kocherlakota, N.** (1996), "The equity premium puzzle: it' s still a puzzle," *Journal of Economic Literature* 34(1), 42 - 71.

[30] **Jegadeesh, N. and Titman, S.** (1993), "Returns to buying winners and selling losers: implications for stock market efficiency," *Journal of Finance* 48, 65 - 91.

[31] **Kreups, D. and Porteus, E.** (1978), "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica* 46, 185 - 200.

[32] **Lucas, R.** (1978), "Asset pricing in an exchange economy," *Econometrica* 46, 1429 - 1445.

[33] **Lustig, H. and Verdelhan, A.** (2006), "The cross-section of foreign currency risk premia and consumption growth risk," *American Economic Review*, forthcoming.

[34] **Mark, N.** (1985), "On time varying risk premiums in the foreign exchange market: an econometric analysis," *Journal of Monetary Economics* 16, 3 - 18.

[35] **Mehra, R. and Prescott, E.** (1985), "The equity premium: a puzzle," *Journal of Monetary Economics* 15, 627 - 636.

[36] **Rubinstein, M.** (1976), "The valuation of uncertain income streams and the pricing of options," *Bell Journal of Economics* 7, 407 - 425.

[37] **Ryder, H. and Heal, G.** (1973), "Optimum growth with intertemporally dependent preferences," *Review of Economics Studies* 40, 1 - 33.

[38] **Samuelson, P.** (1969), "Lifetime portfolio selection by dynamic stochastic programming," *Review of Economics and Statistics* 51, 239 - 246.

[39] **Shiller, R.** (1982), "Consumption, asset markets and macroeconomic fluctuations," *Carnegie Mellon Conference Series on Public Policy* 17, 203 - 238.

[40] **Sundaresan, S.** (1989), "Intertemporally dependent preferences and the volatility of consumption and wealth," *Review of Financial Studies* 2, 73 - 88.

[41] **Verdelhan, A.** (2006), "A habit-based explanation of the exchange rate risk premium," Mimeo.

[42] **Weil, P.** (1989), "The equity premium puzzle and the riskfree rate puzzle," *Journal of Monetary Economics* 24, 401 - 421.

Table 1: Data summary statistics

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

Govern. bonds	US$ real net return on the covered trading of foreign government bonds: Sample Mean (% per year)	Sample S. D. (% per year)	US$ real net return on the uncovered trading of foreign government bonds: Sample Mean (% per year)	Sample S. D. (% per year)	US$ real excess return on the uncovered over the covered trading of foreign govern. bonds: Sample Mean (% per year)	Sample S. D. (% per year)	Annualized Sharp ratio	Quarterly correlation of excess return and US consumption growth
British	1,767	2,224	4,779	10,458	2,972	10,724	0,277	0,050
Canadian	2,273	2,222	2,494	5,515	0,217	4,983	0,044	0,027
German	2,069	2,655	2,944	11,961	0,862	12,320	0,070	0,066
Japanese	2,593	1,809	3,892	13,324	1,274	13,227	0,096	0,141

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

US$ real net return on 90-day Treasury bill: Sample Mean (% per year)	Sample S. D. (% per year)	US$ real net return on S&P500: Sample Mean (% per year)	Sample S. D. (% per year)	US$ real excess return on S&P500 over 90-day T-bill: Sample Mean (% per year)	Sample S. D. (% per year)	Annualized Sharp ratio	Quarterly correlation of excess return and US consumption growth
1,774	1,613	8,952	16,660	7,084	16,288	0,435	0,274

Table 2: Testing CRRA preference [1]

	Equity Premium		Foreign Exchange Risk Premia		Equity and Foreign Exchange Risk Premia	
	$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(R^{SP}_{t+1}-R^{Tb}_{t+1}\right)\right]=0$		$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(R^{U}_{t+1}-R^{C}_{t+1}\right)\right]=0$		$\left\{\begin{array}{l} E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(R^{SP}_{t+1}-R^{Tb}_{t+1}\right)\right]=0 \quad \text{and} \\ E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(R^{U}_{t+1}-R^{C}_{t+1}\right)\right]=0 \end{array}\right.$	
	α	J-statistic	α	J-statistic	α	J-statistic
IS1: Two lags of the US$ real returns on S&P 500, on 90-day Treasury Bill and on uncovered trading of foreign government bonds and three lags of real consumption and real GDP growth rates	95,921 (0.002)	0,253 (0.077)	101,228 (0.000)	0,211 (0.998)	98,731 (0.000)	0,214 (1.000)
	α	J-statistic	α	J-statistic	α	J-statistic
IS2: Current forward premia, the first lag of Fama/French benchmark factors and of the Momentum factor and three lags of real consumption and real GDP growth rates	113,988 (0.005)	0,196 (0.099)	80,920 (0.000)	0,179 (0.998)	77,935 (0.000)	0,200 (0,999)

Notes: P-values in parenthesis

* Coefficient not significant at the 5% significance level.

[1] Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1977:1 to 2004:3.

Table 3: Testing Abel (1990) *catching up with the Joneses* preference [1]

	Equity Premium			Foreign Exchange Risk Premia			Equity and Foreign Exchange Risk Premia		
	$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(\frac{C_t}{C_{t-1}}\right)^{-\kappa(\alpha-1)}\left(R^{SP}_{t+1}-R^{Tb}_{t+1}\right)\right]=0$			$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(\frac{C_t}{C_{t-1}}\right)^{-\kappa(\alpha-1)}\left(R^{U}_{t+1}-R^{C}_{t+1}\right)\right]=0$			$\begin{cases} E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(\frac{C_t}{C_{t-1}}\right)^{-\kappa(\alpha-1)}\left(R^{SP}_{t+1}-R^{Tb}_{t+1}\right)\right]=0 \quad \text{and} \\ E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}\left(\frac{C_t}{C_{t-1}}\right)^{-\kappa(\alpha-1)}\left(R^{U}_{t+1}-R^{C}_{t+1}\right)\right]=0 \end{cases}$		
	α	κ	J-statistic	α	κ	J-statistic	α	κ	J-statistic
IS1: Two lags of the US$ real returns on S&P 500, on 90-day Treasury Bill and on uncovered trading of foreign government bonds and three lags of real consumption and real GDP growth rates	111,313 (0.000)	-0,717 (0.000)	0,120 (0,742)	123,444 (0.000)	-0,857 (0.000)	0,377 (0,997)	108,992 (0.000)	-0,703 (0.000)	0,278 (1,000)
IS2: Current forward premia, the first lag of Fama/French benchmark factors and of the Momentum factor and three lags of real consumption and real GDP growth rates	123,773 (0.000)	-0,535 (0,009)	0,087 (0,923)	198,726 (0.000)	-0,905 (0.000)	0,141 (0,999)	116,503 (0.000)	-0,750 (0.000)	0,376 (0,997)

Notes: P-values in parenthesis

* Coefficient not significant at the 5% significance level.

[1] Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1977:1 to 2004:3.

Table 4: Testing Epstein and Zin (1991) preference [1]

	Equity Premium			Foreign Exchange Risk Premia			Equity and Foreign Exchange Risk Premia		
	$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)}\left(R_{t+1}^{SP}-R_{t+1}^{Tb}\right)\right]=0$			$E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)}\left(R_{t+1}^{U}-R_{t+1}^{C}\right)\right]=0$			$\begin{cases} E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)}\left(R_{t+1}^{SP}-R_{t+1}^{Tb}\right)\right]=0 \text{ and} \\ E_t\left[\left(\frac{C_{t+1}}{C_t}\right)^{\rho(\alpha-1)/(1-\rho)} R_{M,t+1}^{(\rho-\alpha)/(1-\rho)}\left(R_{t+1}^{U}-R_{t+1}^{C}\right)\right]=0 \end{cases}$		
IS1: Two lags of the US$ real returns on S&P 500, on 90-day Treasury Bill and on uncovered trading of foreign government bonds and three lags of real consumption and real GDP growth rates	α	ρ	J-statistic	α	ρ	J-statistic	α	ρ	J-statistic
	96,408	64,639	0,131	106,960	40,854	0,123	110,911	65,186	0,134
	(0.000)	(0,002)	(0,657)	(0.000)	(0.000)	(1,000)	(0.000)	(0.000)	(1,000)
IS2: Current forward premia, the first lag of Fama/French benchmark factors and of the Momentum factor and three lags of real consumption and real GDP growth rates	α	ρ	J-statistic	α	ρ	J-statistic	α	ρ	J-statistic
	473,24	94,170	0,103	389,795	30,688	0,587	400,467	55,975	0,701
	(0,001)	(0.000)	(0,594)	(0.000)	(0.000)	(0,297)	(0.000)	(0.000)	(0,419)

Notes: P-values in parenthesis

* Coefficient not significant at the 5% significance level.B115

[1] Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1977:1 to 2004:3.

Table 5: Testing a modified Campbell and Cochrane (1999) *habit formation* preference [1]

	Equity Premium		Foreign Exchange Risk Premia		Equity and Foreign Exchange Risk Premia	
	$E_t\left[\left(\frac{C_{t+1}S_{t+1}}{C_tS_t}\right)^{-\alpha}\left(\frac{X_{t+1}}{X_t}\right)^{(\alpha-1)}\left(R^{SP}_{t+1}-R^{Tb}_{t+1}\right)\right]=0$		$E_t\left[\left(\frac{C_{t+1}S_{t+1}}{C_tS_t}\right)^{-\alpha}\left(\frac{X_{t+1}}{X_t}\right)^{(\alpha-1)}\left(R^{U}_{t+1}-R^{C}_{t+1}\right)\right]=0$		$\left\{\begin{array}{l} E_t\left[\left(\frac{C_{t+1}S_{t+1}}{C_tS_t}\right)^{-\alpha}\left(\frac{X_{t+1}}{X_t}\right)^{(\alpha-1)}\left(R^{SP}_{t+1}-R^{Tb}_{t+1}\right)\right]=0 \quad \text{and} \\ E_t\left[\left(\frac{C_{t+1}S_{t+1}}{C_tS_t}\right)^{-\alpha}\left(\frac{X_{t+1}}{X_t}\right)^{(\alpha-1)}\left(R^{U}_{t+1}-R^{C}_{t+1}\right)\right]=0 \end{array}\right.$	
IS1: Two lags of the US$ real returns on S&P 500, on 90-day Treasury Bill and on uncovered trading of foreign government bonds and three lags of real consumption and real GDP growth rates	α	J-statistic	α	J-statistic	α	J-statistic
	3,133	0,137	7,319	0,325	7,719	0,361
	(0,031)	(0,676)	(0.000)	(0,997)	(0.000)	(0,999)
IS2: Current forward premia, the first lag of Fama/French benchmark factors and of the Momentum factor and three lags of real consumption and real GDP growth rates	α	J-statistic	α	J-statistic	α	J-statistic
	8,108	0,118	2,891	0,206	5,667	0,209
	(0,001)	(0,545)	(0,029)	(0,998)	(0.000)	(0,999)

Notes: P-values in parenthesis

* Coefficient not significant at the 5% significance level. ** Reject the overidentifying restrictions at the 5% significance level.

[1] Hansens`s (1982) Generalized Method of Moments (GMM) technique is used to test Euler equations and estimate the model parameters over the period from 1977:1 to 2004:3.

Figure 1: Realized vs. fitted excess returns (taking into account the instrument set IS1)

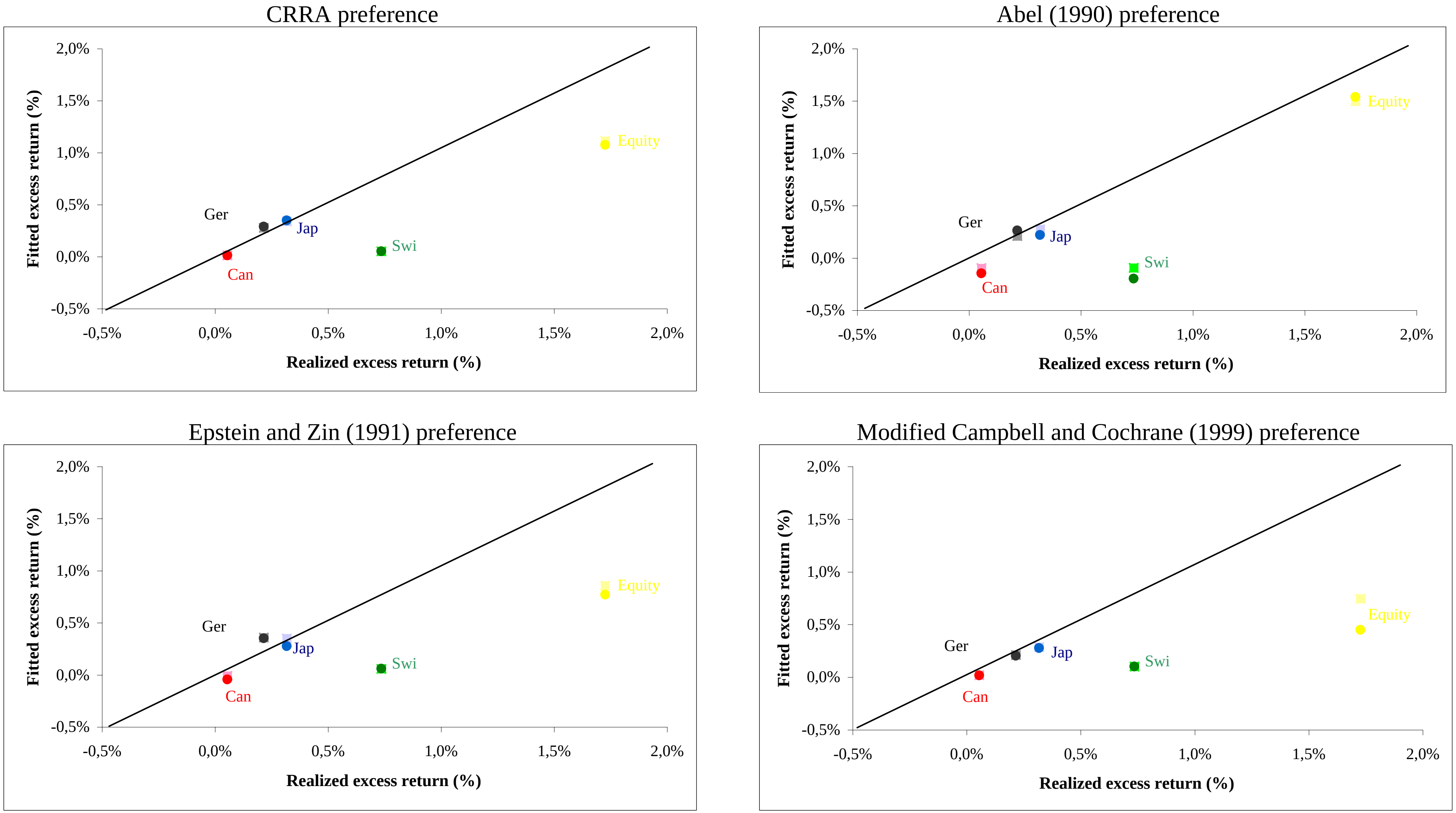
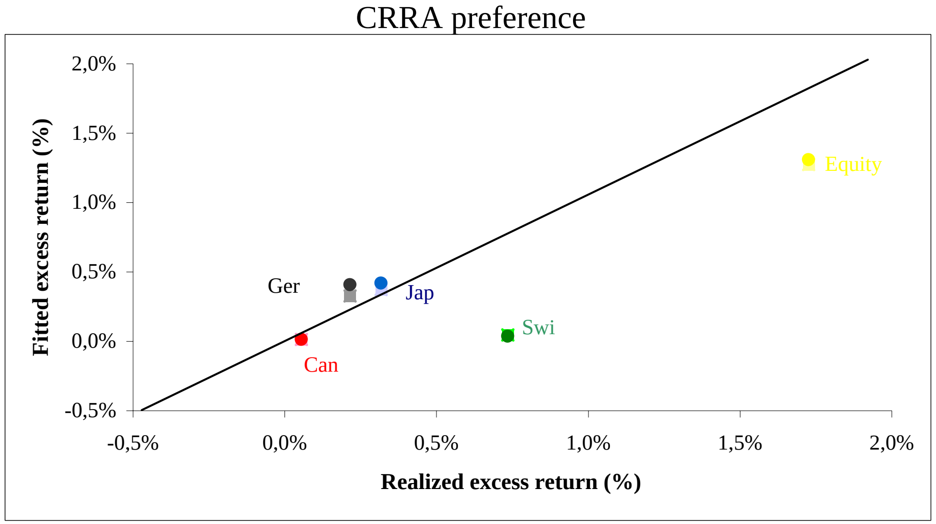
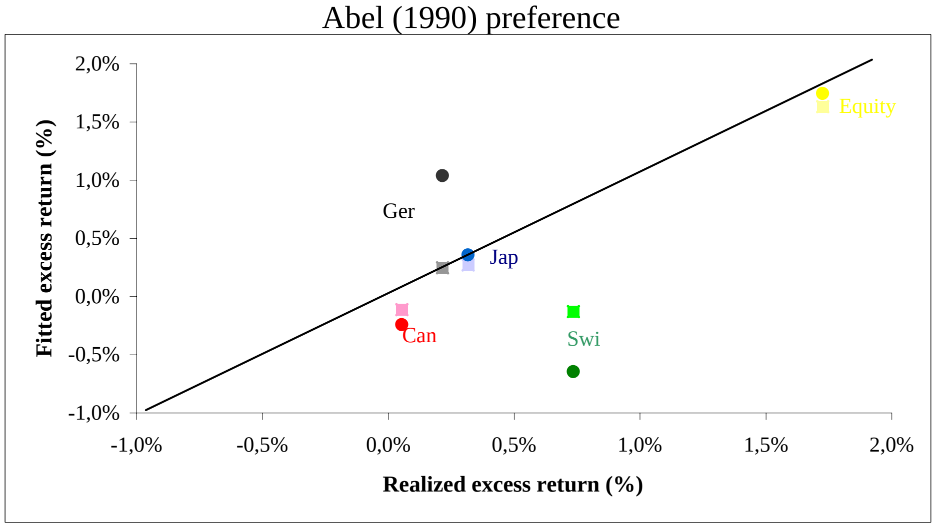
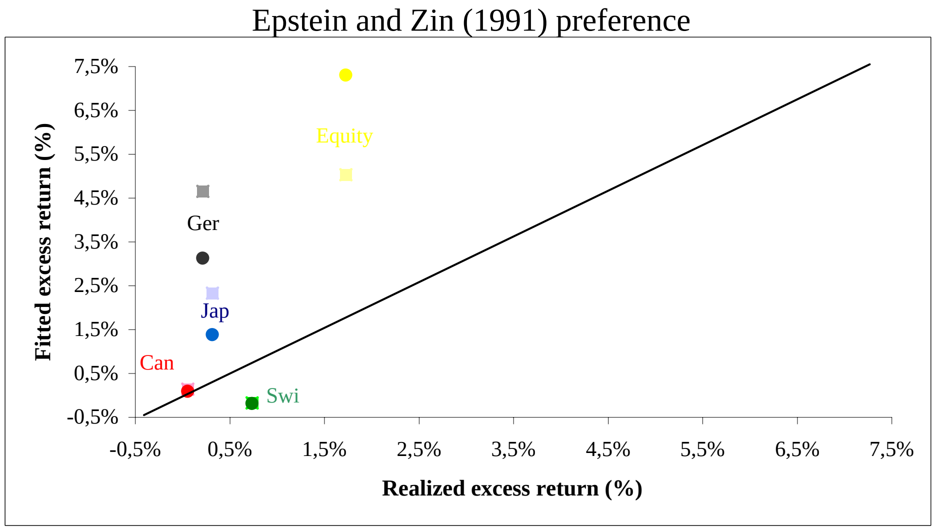
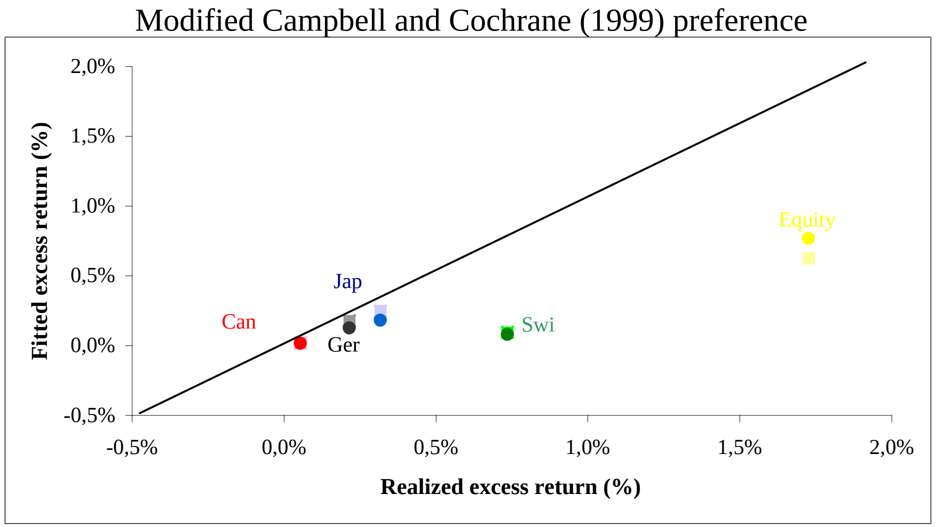


Figure 2: Realized vs. fitted excess returns (taking into account the instrument set IS2)









Notes

[28]First draft November 2006. An earlier version of this paper circulated under the title "Do the Equity and the Foreign Currency Risk Premia Display Common Patterns?"

[29]See the comprehensive surveys by Hodrick (1987) and Engel (1996).

[30]Kocherlakota (1996), for instance, argues that, from an academic economist's perspective, this model is more important than the widely used CAPM.

[31]This statement requires some qualification. The sense in which a representative agent is defined here is not the one proposed by Gorman (1953). There, one asks whether an aggregate function that can be rationalized as the choice of a representative consumer exists for any distribution of wealth. Here, the distribution of wealth is given by a function that maps aggregate wealth to individual wealth. What Constantinides (1982) shows is that, with complete markets, the implied function emulates in the aggregate demand function all the properties implied by the choices of a rational agent.

[32]See e.g. Samuelson (1969), Rubinstein (1976), Lucas (1978), Breeden (1979), Grossman and Shiller (1981) and some other classical papers in macroeconomics and finance.

[33]Here, we are implicitly assuming the absence of short-sale constraints 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).

[34]One must note that disregarding the Euler equation (41), incur the nonidentification of the subjective time discount factor.

[35]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.

[36]The proof of this equivalence as well as the intuition of this assumption can be see in details in Garcia et al. (2006).

[37]The Campbell and Cochrane (1999) model can also accommodate more complex consumption processes, including processes with predictability, conditional heteroskedasticity, and nonnormality.

[38]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 (54).

[39]Proposing this special case of the Garcia et al's framework, we allow for its main

features, except for the reference level to directly affect utility. Although they find strong support for the reference level plays an independent role in the agent's utility function, this is not a consensual evidence which can yet incur in estimates of the subjective time discount factor greater than one.

[40]We recognize that in order to obtain a "perfect" comparison of the behaviors of the equity and the foreign currencey risk premia, the instrument set IS1should include lags of the $US\$$ real returns on the covered trading of foreign government bonds. But, since the covered interest rate parity must hold, in order to deal with the problem of redundant instruments, we use only lags of the real returns on 90-dat T-bill.

[41]See Fama and French (1993) for a complete description of these factor returns.

[42]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.

[43]The results corresponding to this more general framework are not reported in this paper, since they are quite similar to the ones obtained for the catching up with the Joneses model, but are available with the authors.

[44]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.

[45]Once more these results, which are not reported here in order to preserve space and to avoid redundance in some sense, are available with the authors.

[46]As already mentioned, this finding does not hold only in some cases for the original version of the Campbell and Cochrane's model.

CHAPTER 3

Modeling the exchange rate risk premium 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 log-linearizing 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: G12; G15. Keywords: Forward Prmieum Puzzle, GARCH models, Time-varying conditional covariances.

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.[47]

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, ${}_tf_{t+1} - s_t$, is too strongly (negatively) correlated with subsequent changes in the (log of the) exchange rate, $s_{t+1} - s_t$. Domestic currency is expected to appreciate when domestic nominal interest rates exceed foreign 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.[48] 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).[49] 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[50] — henceforth, SDF — 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.

We believe that using this return-based pricing kernel estimated in da Costa et al. (2006), which is compatible with the Forward and the Equity premium behaviors, we can avoid a "joint test", in the sense that an eventual failure of our test may be due only to the log-linearized equation.

The second step then consists in explicitly considering time-varying conditional covariances. To address this issue, we select the more adequate specification to model exchange rate movements, among a representative set of (symmetric and asymmetric) Generalized Autoregressive Conditional Heteroskedasticity models that allow the conditional variance to be a determinant of the mean — i.e. the GARCH-in-mean, or just GARCH-M models.

In the third stage, since we have a time series for this omitted term, we are then able to estimate not only the currency depreciation equation but also its simplified versions with purposes that will be clear later.

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.

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.[51]

Most studies report the FPP through the finding of $\hat{\alpha}_1 < 1$ when running the regression,

$$s_{t+1} - s_t = \alpha_0 + (1 - \alpha_1)({}_t f_{t+1} - s_t) + u_{t+1}, \tag{65}$$

where s_t is the log of time t exchange rate, ${}_t f_{t+1}$ is the log of time t forward exchange rate contract and u_{t+1} is the regression error.[52] Notwithstanding the possible effect of Jensen inequality terms, testing the market-efficiency 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.[53] 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.

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 (65), 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 log-normality 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,

$$1 = \mathbb{E}_t \left(M_{t+1} R^i_{t+1} \right), \tag{66}$$

and

$$0 = \mathbb{E}_t \left[M_{t+1} \left(R^i_{t+1} - R^j_{t+1} \right) \right]. \tag{67}$$

Here, $\mathbb{E}_t(\cdot)$ denotes the conditional expectation given the information available at time t, R^i_{t+1} and R^j_{t+1} represent, respectively, the real gross return on assets i and j at time $t+1$ and M_{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 (66) and (67). 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_{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 SDF, M^*_{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 SDF, M_{t+1}, may thus be written as $M_{t+1} = M^*_{t+1} + \nu_{t+1}$ for some ν_{t+1} with $\mathbb{E}_t\left[\nu_{t+1} R^i_{t+1}\right] = 0\ \forall i$.

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 (67) can be useful to analyze the FPP.[54]

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

$$R^C_{t+1} = \frac{{}_tF_{t+1}(1+i^*_t)P_t}{S_t P_{t+1}} \qquad \text{and} \qquad R^U_{t+1} = \frac{S_{t+1}(1+i^*_t)P_t}{S_t P_{t+1}}, \tag{68}$$

where ${}_tF_{t+1}$ and S_t are the forward and spot prices of foreign currency in terms of domestic currency, P_t is the dollar price level and i^*_t represents nominal net return on a foreign asset in terms of the foreign investor's preferences.[55]

2.2.2. Currency depreciation theoretical framework

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

Assumption 1: The Asset Pricing Equation $\mathbb{E}_t\left(M_{t+1}R^i_{t+1}\right) = 1$ 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 $M_{t+1}R^i_{t+1}$ 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 $M_{t+1}R^i_{t+1}$ is lognormal.

Taking logs in both sides of (16) and further applying a Taylor expansion yields, for every asset i in the economy, we have

$$r^i_{t+1} = -m_{t+1} - \frac{1}{2}(\sigma^i_t)^2 + \varepsilon^i_{t+1}, \tag{69}$$

where the variables in small letters are the logs of the variables in capital letters, $\varepsilon^i_{t+1} = \ln\left(M_t R^i_{t+1}\right) - \mathbb{E}_t\{\ln\left(M_{t+1}R^i_{t+1}\right)\}$ denotes the innovation in predicting $\ln\left(M_t R^i_t\right)$, $(\sigma^i_t)^2 \equiv \mathbb{V}_t(\ln M_{t+1} + \ln R^i_{t+1})$ and $\mathbb{V}_t(\cdot)$ denotes the conditional variance given the available information at time t. From (69) one verifies that asset returns are decomposed in three terms. The first one is the logarithm of the SDF, m_{t+1}, which is common to all returns and is a random variable. The second one, $(\sigma^i_t)^2$, is idiosyncratic and, given past information, also deterministic, but not necessarily constant. The third one is ε^i_{t+1}. 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_t.

Because equation (69) 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

$$s_{t+1} - s_t = \frac{1}{2}[(\sigma^C_t)^2 - (\sigma^U_t)^2] + ({}_t f_{t+1} - s_t) + \varepsilon^U_{t+1} - \varepsilon^C_{t+1}, \tag{70}$$

where $\sigma_t^C \equiv \mathbb{V}_t(\ln M_{t+1} + \ln R_{t+1}^C)$ and $\sigma_t^U \equiv \mathbb{V}_t(\ln M_{t+1} + \ln R_{t+1}^U)$.

Consequently,

$$s_{t+1} - s_t = -\left[\frac{1}{2}\mathbb{V}_t\left(s_{t+1}\right) + cov_t\left(s_{t+1}, m_{t+1} - \pi_{t+1}\right)\right] + \left({}_tf_{t+1} - s_t\right) + \varepsilon_{t+1}, \qquad (71)$$

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

This equation (71) is useful in drawing implications for a SDF that accounts for the FPP.

Taking the conditional expectation on (71) and on (65)[56], it is straightforward to see that

$$\mathbb{E}_t(s_{t+1} - s_t) = -\left[\frac{1}{2}\mathbb{V}_t\left(s_{t+1}\right) + cov_t\left(s_{t+1}, m_{t+1} - \pi_{t+1}\right)\right] + \left({}_tf_{t+1} - s_t\right)$$

$$\mathbb{E}_t^*(s_{t+1} - s_t) = \alpha_0 + \alpha_1({}_tf_{t+1} - s_t)$$

and consequently,

$$({}_tf_{t+1} - s_t)(1 - \alpha_1) = \alpha_0 + \mathbb{E}_t^*\left[\frac{1}{2}\mathbb{V}_t\left(s_{t+1}\right) + cov_t\left(s_{t+1}, m_{t+1} - \pi_{t+1}\right)\right]. \qquad (72)$$

Denoting by $\xi_t(m_{t+1}) \equiv \mathbb{E}_t^*\left[\frac{1}{2}\mathbb{V}_t\left(s_{t+1}\right) + cov_t\left(s_{t+1}, m_{t+1} - \pi_{t+1}\right)\right]$ and by $\mu_i(x) \equiv$ ith 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 α_1, it is necessary and sufficient that the SDF satisfies

$$\begin{cases} \mu_i(\xi_t(m_{t+1})) > \mu_i({}_tf_{t+1} - s_t), & \text{for } i \text{ even} \\ \begin{array}{c} |\mu_i(\xi_t(m_{t+1}))| > |\mu_i({}_tf_{t+1} - s_t)| \text{ and} \\ \text{both moments with the same sign,} \end{array} & \text{for } i \text{ odd,} \end{cases} \qquad (73)$$

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 (71) is also extremely useful in deriving empirical tests of FPP.[57]

First, one should note that the conventional regression (65) is a particular case of (71), where the term $\frac{1}{2}\mathbb{V}_t(s_{t+1}) + cov_t(s_{t+1}, m_{t+1} - \pi_{t+1})$ is assumed to be time-invariant. Based on (71), i.e. in a lognormal world, it is apparent that the disappointing empirical findings reported when regression (65) 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 $\mathbb{V}_t(s_{t+1})$ 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 (65) is not responsible for finding $\hat{\alpha}_1 < 0$. Therefore, all the action must come from the term $cov_t(s_{t+1}, m_{t+1} - \pi_{t+1})$. 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,[58] 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

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

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

Because every asset return R^i_{t+1} contains "a piece" of M_{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 $N \to \infty$ or $N, T \to \infty$, 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).

Multifactor Models Factor models summarize the systematic variation of the N elements of the vector $\mathbf{R}_t = \left(R^1_t, R^2_t, \ ..., R^N_t\right)'$ using a reduced number of K factors, $K < N$. Consider a K-factor model in R^i_t:

$$
R^i_t = a_i + \sum_{k=1}^{K} \beta_{i,k} f_{k,t} + \eta_{it} \tag{74}
$$

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

$$
\text{plim} \frac{1}{N} \sum_{i=1}^{N} \eta_{i,t} = 0.
$$

Denote by $\Sigma_{\mathbf{r}} = \mathbb{E}\left(\mathbf{R}_t\mathbf{R}_t'\right) - \mathbb{E}\left(\mathbf{R}_t\right)\mathbb{E}\left(\mathbf{R}_t'\right)$ the variance-covariance matrix of returns. The first principal component of the elements of $\mathbf{R}_t$ is a linear combination $\theta'\mathbf{R}_t$ with maximal variance subject to the normalization that θ has unit norm, i.e., $\theta'\theta = 1$. 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 $\mathbf{R}_t$ are consistent estimates of the $f_{k,t}$'s. Factor loadings can be estimated consistently by simple OLS regressions of the form (74).

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 $\widetilde{M_t^*}$, 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_t^i in (74), one can write the respective expected-beta return expression

$$\mathbb{E}(R^i) = \gamma + \sum_{k=1}^{K} \beta_{i,k}\lambda_k, i = 1, 2, ..., N$$

where λ_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 λ coefficients directly by

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

and consequently to estimate only γ via a cross-sectional regression[59].

Given these coefficients, can easily get an estimate of M_t^*. The relation between (λ, γ) and (a, b), given by

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

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

$$\widetilde{M_t^*} = a + \sum_{k=1}^{K} b_k f_{k,t} \text{ and } \mathbb{E}(\widetilde{M_t^*} R^i) = 1, i = 1, 2, ..., N$$

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_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_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 (70) and (71) 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 $\mathbb{V}_t(\ln M_{t+1} + \ln R_{t+1}^C)$ and $\mathbb{V}_t(\ln M_{t+1} + \ln R_{t+1}^U)$ 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 (70) due to its analytical simplicity, although we still intend to analyse the individual role played by the $cov_t\left(s_{t+1}, m_{t+1} - \pi_{t+1}\right)$ term in the equation (71) 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

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

where the expectations of the process $m_{t+1} + r_{t+1}^U$ 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 $m_{t+1}+r_{t+1}^i$ process is that the literature on conditional heteroskedasticity models is so extensive,[60] 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,[61] 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.

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			
ARCH(1)-M and	ARCH(2)-M		
GARCH(1,1)-M	GARCH(2,1)-M	GARCH(1,2)-M and	GARCH(2,2)-M

Asymmetric models			
TARCH(1)-M	TARCH(2)-M		
TARCH(1,1)-M	TARCH(2,1)-M	TARCH(1,2)-M and	TARCH(2,2)-M
EARCH(1)-M	EGARCH(1,1)-M and	EGARCH(2,1)-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
$m_{t+1} + r^U_{t+1} = \phi(\sigma^U_t)^2 + \varepsilon^i_{t+1}$

Variance equation	
Symmetric models:	
ARCH (q) model	$(\sigma^U_t)^2 = \varpi + \sum_{j=1}^{q} \delta_q(\varepsilon^U_{t+1-q})^2$
GARCH (p,q) model	$(\sigma^U_t)^2 = \varpi + \sum_{i=1}^{p} \eta_i(\sigma^U_{t-i})^2 + \sum_{j=1}^{q} \delta_q(\varepsilon^U_{t+1-q})^2$
Asymmetric models:	
TARCH (p,q)	$(\sigma^U_t)^2 = \varpi + \sum_{i=1}^{p} \eta_i(\sigma^U_{t-i})^2 + \gamma d_t(\varepsilon^U_t)^2 + \sum_{j=1}^{q} \delta_q(\varepsilon^U_{t+1-q})^2$
EGARCH (p,1)	$\ln[(\sigma^U_t)^2] = \varpi + \sum_{i=1}^{p} \theta_i \ln[(\sigma^U_{t-i})^2] + \gamma_1 \frac{\varepsilon^U_t}{\sigma^U_{t-i}} + \lambda_1 \left\| \frac{\varepsilon^U_t}{\sigma^U_{t-i}} \right\|$

3.3. Adjusted regressions

In this subsection we describe the last step. Since we have estimated the omitted term $[(\sigma^C_t)^2 - (\sigma^U_t)^2]$, we can analyse its relevance in accommodating the empirical evidences related to the FPP.

First we run the conventional regression (65) which works as our benchmark. This regression can be considered as a special case of (70), where the omitted term assumes a null value.

Next, we propose running the regression

$$s_{t+1} - s_t = (0.5 - \psi)\varphi_t + (1 - \alpha_1)({}_t f_{t+1} - s_t) + \varepsilon_{t+1}, \tag{75}$$

where $\varphi_t \equiv [(\sigma_t^C)^2 - (\sigma_t^U)^2]$. According to (70), 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 (75). When we assume that $\hat{\psi} = 0$, we are able to evidence the power of the forward premium to predict the term $s_{t+1} - s_t - 0.5\varphi_t$, where the null hypothesis consists in testing $\hat{\alpha}_1 = 0$ in regression

$$s_{t+1} - s_t - 0.5\varphi_t = (1 - \alpha_1)({}_t f_{t+1} - s_t) + \varepsilon_{t+1}. \tag{76}$$

Finally, assuming $\hat{\alpha}_1 = 0$ we run the regression

$$s_{t+1} -_t f_{t+1} = (0.5 - \psi)\varphi_t + \varepsilon_{t+1}, \tag{77}$$

which enables us to evidence the power of the omitted term to explain the term $s_{t+1} -_t f_{t+1}$. 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[62]. 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_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[63], 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_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_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_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 $\widetilde{M}_t$ 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 *ii)* 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 $\widetilde{M}_t$ 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 φ_t.

According to our results reported in table 2, where we show the conditional variance models chosen for both process $m_{t+1} + r^U_{t+1}$ and $m_{t+1} + r^C_{t+1}$ 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 (75), 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 (76), except for the Swiss case, the inclusion of the $-\frac{1}{2}[(\sigma^C_t)^2 - (\sigma^U_t)^2]$ term in the left hand side of the conventional regression seems to be useful in reducing considerably the value of the coefficient α_1, but once more not enough to provide the "correct" forward predictability power.

Finally, according to the results of the adjusted regression (77), even for the Japanese case where the coefficient ψ 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 $m_{t+1} + r_{t+1}^{U}$ and $m_{t+1} + r_{t+1}^{C}$ 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[64] and *ii)* tests of stationarity and cointegration of the appropriate variables.

Finally we still intend to analyse the individual role played by the $cov_t\left(s_{t+1}, m_{t+1} - \pi_{t+1}\right)$ term in the equation (71) using a multivariate GARCH(p,q)-M model, following the approaches proposed by Goeij and Marquering (2002) and Bollerslev et al. (1992).

5. Conclusion

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.

[8] **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.

[9] **Campbell, J. Y. , Lo, A. W. and MacKinlay, A. C.** (1997), "The econometrics of financial markets", Princeton University Press.

[10] **Chamberlain, Gary, and Rothschild, Michael** (1983). “Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets,” *Econometrica*, vol. 51(5), pp. 1281-1304.

[11] **Cochrane, J.H.** (2001),"Asset Pricing", Princeton Univerity Press.

[12] **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.

[13] **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.

[14] **Connor, G. and Korajczyk, R.** (1993), "A test for the number of factors in am approximate factor structure", *Journal of Finance* 48, 1263 - 1291.

[15] **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.

[16] **Cumby, R. E.** (1988) "Is it risk? Explaining deviations from uncovered interest parity", *Journal of Monetary Economics* 22, 279 - 299.

[17] **Engel, C.** (1996), "The forward discount anomaly and the risk premium: a survey of recent evidence", *Journal of Empirical Finance* 3, 123 - 192.

[18] **Engle, R.F.** (1982) “Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation”, *Econometrica* 50 (3), 987-1008.

[19] **Engle, R.F. and Bollerslev, T.** (1986) “Modelling the persistence of conditional variances”, *Econometric Reviews* 5, 1-50.

[20] **Engle, R.F., T. Ito, and Lin, W. L.** (1990) “Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market,” *Econometrica* 58 (3), 525-542.

[21] **Engle, R.F., Lilien, D. M. and Robins, R. P.** (1987) “Estimating time varying risk premia in the term structure: the ARCH-M model”, *Econometrica* 55 (2), 391-407.

[22] **Fama, E. F.** (1975), "Short-term interest rates as predictors of inflation", *American Economic Review*, 65, 269 - 282.

[23] **Fama, E. F.** (1984), "Term premiums in bond returns", *Journal of Financial Economics*, 13, 529 - 546.

[24] **Frankel, J. A.** (1979), "The diversifiability of exchange risk", *Journal of International Economics* 9, 379-393.

[25] **Froot, K. A.** (1990), "Short rates and expected asset returns," NBER Working paper n^o 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 sterling exchange rate risk premia", *Journal of International Money and Finance* 25, 71 - 92.

[28] **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.

[29] **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.

[30] **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.

[31] **Hansen, L. and Singleton, K.** (1982) "Generalized instrumental variables estimation of nonlinear expectations models," *Econometrica* 50(5), pp. 1269-1286.

[32] **Hansen, L. and Singleton, K.** (1983), "Stochastic consumption, risk aversion and the temporal behavior of asset returns". *Journal of Political Economy* 91(2), 249-265.

[33] **Hansen, L. and Singleton, K.** (1984) "Erratum of the article: 'Generalized instrumental variables estimation of nonlinear expectations models,' " *Econometrica* 52(1), pp. 267-268.

[34] **Harrison, J. M. and Kreps, D. M.** (1979), "Martingales and arbitrage in multiperiod securities markets," *Journal of Economic Theory* 20, 249 - 268.

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

[36] **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.

[37] **Huang, R. D.** (1989), "An analysis of intertemporal pricing for forward exchange contracts", *Journal of Finance* 44, 183 - 194.

[38] **Lehmann, B. and Modest, D.** (1988), "The empirical foundations of the arbitrage pricing theory", *Journal of Financial Economics*, 21, 213-254.

[39] **Lewis, K. K.** (1990), "The behavior of Eurocurrency returns across different holding periods and regimes", *Journal of Finance* 45, 1211 - 1236.

[40] **Lucas, R. E.** (1982), "Interest rates and currency prices in a two-country world," *Journal of Monetary Economics* 10, 335 - 360.

[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.

Appendix

A. Tables and figures

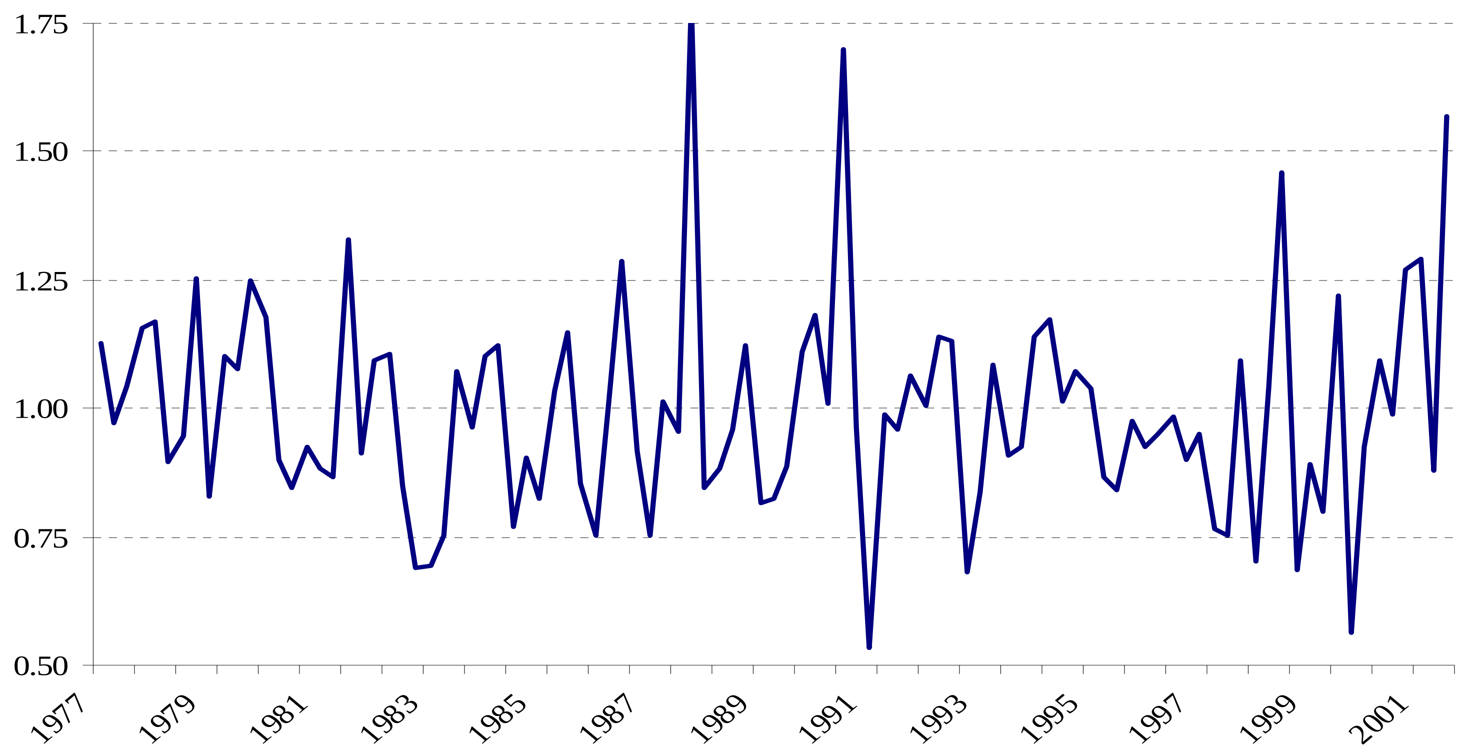
Table 1: Data summary statistics

International quarterly data: observations from 1977:I to 2001:III

	$US\$$ real return on the covered trading of foreign government bonds		$US\$$ real return on the uncovered trading of foreign government bonds		$US\$$ real excess return on the uncovered over the covered trading of foreign government bonds	
Government Bonds	Sample mean (% per year)	Sample stand. deviat. (% per year)	Sample mean (% per year)	Sample stand. deviat. (% per year)	Sample mean (% per year)	Sample stand. deviat. (% 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

Notes: Means and standard deviations are given in annualized percentage points. To annualize the raw quarterly numbers, means are multiplied by 400 while standard deviations are multiplied by 200.

Figure 1: SDF estimate using multifactor model (1977:1 - 2001:3): $\widetilde{M}_t$



	Mean	Stan. dev.	Max.	Min.
$\widetilde{M}_t$	0.995	0.210	1.797	0.532

Table 2: Models of conditional variance

Canadian government bonds

Conditional variance	Model	Mean equation	Variance equation			
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\eta}_1$	$\hat{\gamma}$	$\hat{\delta}_1$
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R^{C}_{t+1})]$	TARCH (1,1) - M^{1,2}	-0.429	0.024	0.501	-0.197	0.057
		(0.400)*	(0.000)	(0.000)	(0.076)*	(0.571)*
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\delta}_1$		
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R^{U}_{t+1})]$	ARCH(1) - M^{1,2}	0.333	0.043	-0.094		
		(0.169)*	(0.000)	(0,000)		

German government bonds

Conditional variance	Model	Mean equation	Variance equation		
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\delta}_1$	
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R^{C}_{t+1})]$	ARCH(1) - M^{2}	-0.281	0.045	-0.091	
		(0.542)*	(0.000)	(0.006)	
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\eta}_1$	$\hat{\delta}_1$
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R^{U}_{t+1})]$	GARCH(1,1) - M^{2}	-0.352	0.085	-0.508	-0.080
		(0.421)*	(0.001)	(0.262)*	(0.014)

Japanese government bonds

Conditional variance	Model	Mean equation	Variance equation				
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\gamma}_1$	$\hat{\lambda}_1$
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R^{C}_{t+1})]$	EGARCH(2,1) - M^{1}	-0.243	-10.145	-1.237	-0.918	-0,019	-0.317
		(0,681)*	(0.000)	(0.000)	(0.000)	(0.810)*	(0.028)
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\delta}_1$			
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R^{U}_{t+1})]$	ARCH(1) - M^{2}	-0.290	0.050	-0.070			
		(0.525)*	(0.000)	(0.032)			

Table 2 (continued)

Swiss government bonds

Conditional variance	Model	Mean equation	Variance equation			
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\eta}_1$	$\hat{\delta}_1$	
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R_{t+1}^C)]$	GARCH(1,1) - $M^{1,2}$	0.053	0.020	0.646	-0.111	
		$(0.915)^*$	$(0.190)^*$	$(0.052)^*$	(0.001)	
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\eta}_1$	$\hat{\delta}_1$	$\hat{\delta}_2$
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R_{t+1}^U)]$	GARCH(1,2) - 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

Conditional variance	Model	Mean equation	Variance equation				
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\eta}_1$		$\hat{\gamma}$	$\hat{\delta}_1$
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R_{t+1}^C)]$	TARCH(1,1) - $M^{1,2}$	-0.427	0.022	0.533		0.202	0.057
		$(0.381)^*$	$(0.194)^*$	$(0.168)^*$		(0.000)	(0.000)
		$\hat{\phi}$	$\hat{\varpi}$	$\hat{\eta}_1$	$\hat{\eta}_2$	$\hat{\gamma}$	$\hat{\delta}_1$
$\sigma_t^2[\ln(\widetilde{M_{t+1}}R_{t+1}^U)]$	TARCH(2,1) - M^1	-0.347	0.020	1.311	-0.676	-0.113	0.014
		$(0.459)^*$	(0.036)	(0.000)	(0.015)	$(0.061)^*$	$(0.716)^*$

Notes: * Not significant estimates at the 5% significance level. P - values in the parenthesis. [1] Akaike criteria. [2] Schwarz criteria

Table 3: Conventional and adjusted regressions

	Foreign government bonds									
	Canadian		German		Japanese		Swiss		British	
Conventional regression:	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_0$	$\hat{\alpha}_1$
$s_{t+1} - s_t = \alpha_0 +$	0.003	0.472	0.004	1.713	0.028	3.257	-0.002	0.748	-0.006	1.649
	(0.173)*	(0.072)*	(0.520)*	(0.007)	(0.003)	(0.000)	(0.814)*	(0.000)	(0.397)*	(0.118)*
$+(1-\alpha_1)({}_tf_{t+1} - s_t) + \varepsilon_{t+1}$										
Adjusted regressions:										
	$\hat{\psi}$	$\hat{\alpha}_1$	$\hat{\psi}$	$\hat{\alpha}_1$	$\hat{\psi}$	$\hat{\alpha}_1$	$\hat{\psi}$	$\hat{\alpha}_1$	$\hat{\psi}$	$\hat{\alpha}_1$
i) $s_{t+1} - s_t = (0.5-\psi)\varphi_t +$	0.113	0.282	0.712	1.635	0.574	1.467	0.471	0.753	1.583	1.824
	(0.587)*	(0.255)*	(0.143)*	(0.009)	(0.106)*	(0.017)	(0.060)*	(0.000)	(0.002)	(0.017)
$+(1-\alpha_1)({}_tf_{t+1} - s_t) + \varepsilon_{t+1}$										
	$\hat{\alpha}_1$		$\hat{\alpha}_1$		$\hat{\alpha}_1$		$\hat{\alpha}_1$		$\hat{\alpha}_1$	
ii) $s_{t+1} - s_t - 0.5\varphi_t =$	0.261		1.205		1.189		0.711		1.233	
	(0.284)*		(0.029)		(0.044)		(0.001)		(0.183)*	
$=(1-\alpha_1)({}_tf_{t+1} - s_t) + \varepsilon_{t+1}$										
	$\hat{\psi}$		$\hat{\psi}$		$\hat{\psi}$		$\hat{\psi}$		$\hat{\psi}$	
iii) $s_{t+1} -_t f_{t+1} = (0.5-\psi)\varphi_t + \varepsilon_{t+1}$	0.076		-0.104		0.332		0.039		0.068	
	(0.712)*		(0.828)*		(0.340)*		(0.911)*		(0.900)*	

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.

Notes

[47]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.

[48]We do not claim originality in this characterization.

[49]Additional references are Hansen and Jagganathan (1991) and Bai (2005).

[50]In what follows, we use the terms pricing kernel and stochastic discount factor interchangeably.

[51]See the comprehensive surveys by Hodrick (1987) and Engel (1996).

[52]In what follows we use capital letters to denote variables in levels and small letters to denote the logs of these variables.

[53]As we shall see later, weaker, though still restrictive, conditions on the behavior of risk premium may be compatible with $\alpha_1 = 0$.

[54]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).

[55]Here, we are implicitly assuming that this is a frictionless economy.

[56]Here, $\mathbb{E}_t^*(\cdot)$ denotes the conditional expectation given the forward premium, ${}_tf_{t+1}-s_t$.

[57]ince, according to the literature, $(s_{t+1}-s_t)\tilde{}I(0)$ and $s_t\tilde{}I(1)$, one must be careful when defining the $\mathbb{V}_t\left(s_{t+1}\right)$ and $cov_t\left(s_{t+1}, m_{t+1}-\pi_{t+1}\right)$ terms. Otherwise one may end up with an unbalanced regression.

[58]For more details about the heteroskedasticity of asset returns, see Bollerslev et al. (1988), Engle et al. (1990) and Engle and Marcucci (2005).

[59]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 γ, one set it equal to the real return of the risk free rate.

[60]See Bollerslev et al. (1992) for an excellent survey about conditional variance models.

[61]See, for exmaple, West and Cho (1995) and Malmsten and Terasvirta (2004).

[62]Chicago Mercantile Exchange pioneered the development of financial futures with the launch of currency futures, the world's first financial futures contracts, in 1972.

[63]Data on the return on real estate are measured using the return of all publicly traded REITs – Real-Estate Investment Trusts.

[64]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 sample sizes and very persistent autocorrelation in the forward premium.