Bonds, Currencies and Expectational Errors

Markus Sihvonen*

September 14, 2018

Abstract

I propose a simple no-arbitrage model for explaining bond and currency returns that is consistent with the systematic expectational errors documented in surveys. The implied term premium is on average positive but time-varying due to expectational errors. The model matches the violations of uncovered interest rate parity and creates the observed downward sloping term structure of carry trade returns because underpriced currencies tend to have overpriced long-term bonds. It also explains other currency and bond anomalies such as delayed overshooting, the currency variance and persistence puzzles and the puzzle of excess volatility of long-term yields. It matches survey evidence in that forecasters (i) underweight the importance of recent interest rate shocks on future interest rates, (ii) underestimate the future strength of high interest rate currencies and (iii) overstate the impact of an upward sloping yield curve on future interest rates.

1 Introduction

There are two main alternatives to explaining the predictability of bond and currency returns. First the predictability can be due to variation in required compensation for risk. Recently this approach has been followed by, for example, Bansal and Shaliastovich (2012), Ermolov (2014) and Zviadadze (2016), who propose joint explanations for bond and currency dynamics.

*Bank of Finland, Research Unit. I thank Anna Cieslak, Matti Suominen and seminar participants at the Bank of Finland Research Seminar and HEC Paris Finance PhD Workshop for useful comments. I thank Thomas Philippon for support.

The second alternative is that the predictability represents expectational errors: increased expected returns are related to underpricings in the relevant assets. The view that expectations are not fully rational is supported by the expectational errors documented in survey data (Bacchetta et al. (2009), Greenwood and Shleifer (2014)).[1]

This paper pursues the latter the option. The idea that currency returns are driven by mispricings has been explored by Froot and Frankel (1989), McCallum (1994), Gourinchas and Tornell (2004) and Burnside et al. (2011). Similarly, the effects of belief distortions on interest rates have been studied by, for example, Froot (1989), Xiong and Yan (2010), Hong and Sraer (2013), Piazzesi et al. (2015) and Cieslak (2016). However, to my knowledge this is the first paper to offer a joint explanation for variation in bond and currency returns based on expectational errors. The model of this paper generates the following empirical facts:

1. Currency puzzles
    - Forward premium puzzle and the violations of uncovered interest rate parity (e.g, Fama (1984))
    - Delayed overshooting (e.g., Eichenbaum and Evans (1995))
    - The persistence and volatility puzzles of currencies (e.g., Backus et al. (1993), Moore and Roche (2002))
2. Bond facts/puzzles (⋆)
    - Yield curve is upward sloping on average (⋆) (e.g., Fama and Bliss (1987))
    - Time-variation in expected bond returns (⋆) (e.g., Fama and Bliss (1987), Cochrane and Piazzesi (2005))
    - The puzzle of excess volatility of long-term yields (⋆) (e.g., Shiller (1979))
3. Joint bond-currency puzzles (⋆)

[1] There are some models that do not fall strictly into either category. For example Bacchetta and Wincoop (2010) assume agents that update their portfolios infrequently. These types of limitations can represent either cognitive or more concrete frictions.

- Term structure of expected carry trade returns is downward sloping (⋆) (Lustig et al. (2017))[2]

4. Expectational errors in forecasts

   - Short-term interest rate expectations underreact to recent interest rate shocks (e.g., Gourinchas and Tornell (2004))
   - High interest rate currencies remain stronger than predicted (e.g., Froot and Frankel (1989), Bacchetta et al. (2009))
   - Forecasters overpredict the impact of an upward sloping yield curve on future interest rates (⋆) (e.g., Bacchetta et al. (2009))

The starred points represent empirical facts left unexplained in the analysis of Gourinchas and Tornell (2004). Overall the logic of the model follows that in Gourinchas and Tornell (2004), who find evidence that forecasters overweight the importance of transitory interest rate shocks. This implies overly persistent expectations that underreact to interest rate changes. The authors do not offer a full explanation for such biases, though Ilut (2012) argues that similar effects can be created using ambiguity averse preferences. However, the key idea is that there is still some structural consistency in these interest rate expectations and the pricing of other assets.[3]

Due to higher yields, high interest rate currencies are more valuable than low interest rate currencies. At the same time these currencies are undervalued as forecasters do not believe the high interest rate environment will persist. As in Gourinchas and Tornell (2004), this explains the violations of uncovered interest parity.

Because interest rates are expected to decline, the slope of the yield curve is low. These high relative prices of long-term bonds are partly justified. However, these bonds are overpriced as short-term rates tend to remain high for a longer period than predicted. As a novel prediction, the time-variation in bond risk premia is therefore also primarily caused by expectational errors. Evidence for such mispricing is given, for example, by Bacchetta et al. (2009).

[2]See also Zviadadze (2016) and Chinn and Meredith (2005). Note that I use the term expected returns to refer to objective rather than subjective expectations. This is consistent with the vocabulary of Greenwood and Shleifer (2014).

[3]Clear structural consistency especially among different maturity bonds is imposed by no-arbitrage bounds.

Lustig et al. (2017) find that the term structure of carry trade risk premia is downward sloping. Uncovered interest rate parity holds relatively well for the returns of long-term bonds. This is because high interest rate currencies offer small term premia. The authors argue that this finding is inconsistent with most existing risk-based currency models.[4]

The model in this paper offers a new explanation for this type of joint variation in bond and currency premia. As explained before, a high interest rate currency tends to be underpriced but at the same time long-term bonds of the same currency tend to be overpriced. The term structure of expected carry trade returns is downward sloping because these under- and overpricings partly offset each other.

The key results of this paper can be explained under the assumption that currencies are subject to similar perceived risk premia.[5] Denote the log short-term interest rate differential, also known as forward premium, between the foreign and home country by $x_t \equiv i_t^* - i_t$ and the log FX rate by s_t. The logarithmic perceived uncovered interest rate parity condition is

$$\mathbb{E}_t^S[s_{t+1}] - s_t + x_t = 0, \tag{1}$$

where S denotes the subjective probability measure of the agents. Roughly, this states that the perceived expected return from borrowing in the home currency and investing in the foreign currency is zero. For simplicity assume a stationary nominal exchange rate and a long-run expected log exchange rate of 0 (e.g. due to symmetric countries).[6] From this one can solve

$$s_t = \sum_{i=0}^{\infty} \mathbb{E}_t^S[x_{t+i}]. \tag{2}$$

Given persistent interest rates, the foreign currency is strong after shocks that raise foreign interest rates above home interest rates: $x_t > 0$. The violations of uncovered interest parity are due to the fact that now the interest rate differential tends to remain higher than expected $\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}] > 0$. This is because the forecasters are slow at increasing their interest rate forecasts after the positive interest rate shocks. On the other hand, this implies that $\mathbb{E}_t[s_{t+1}] - \mathbb{E}_t^S[s_{t+1}] > 0$. Effectively the foreign currency tends to remain strong for a longer period than expected.

[4] Moreover, it is inconsistent e.g. with the behavioral currency model of Burnside et al. (2011). It can be shown that the model implies very similar overshooting in short- and long-term rates and therefore fails to satisfy the long bond return parity condition.

[5] Given the symmetric model of the paper, this case emerges when $\varphi_1 = 0$.

[6] I discuss the role of the permanent component of the FX rate later.

The relative log price of a zero coupon bond of maturity n is

$$q_t^*(n) - q_t(n) = -\sum_{i=0}^{n-1} \mathbb{E}_t^S [x_{t+i}]. \tag{3}$$

The price of the foreign bond is relatively low and interest rate high when the foreign short-term interest rate is relatively high $x_t > 0$. At the same time the interest rate tends to remain higher than expected $\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}] > 0$ and therefore $\mathbb{E}_t[q_{t+1}^*(n-1) - q_{t+1}(n-1)] - \mathbb{E}_t^S[q_{t+1}^*(n-1) - q_{t+1}(n-1)] < 0$. The misestimation of the interest rate process therefore creates variation in bond risk premia as high interest rate currencies have relatively overpriced long-term bonds.

Why does this type of model explain the joint behaviour of bonds and currencies? When $x_t > 0$ foreign currency short-term securities have high returns due to violations of the uncovered interest parity. At the same time the long-term bond of the same currency is relatively overpriced and yields low returns. The effect is stronger the longer the maturity of the bonds. One can see that these effects partly offset each other so that a strategy that buys a long-term bond of the foreign currency and sells a similar bond of the home currency yields small returns. This explains why the term structure of expected carry trade returns is downward sloping.

In the main model of this paper all time-variation in expected returns is due to biased beliefs. More specifically, under the agents' biased subjective probability measure S, the expected excess return of bond or currency i is constant

$$\mathbb{E}_t^S [r_{t+1}^{ex}(i)] = constant(i). \tag{4}$$

Note that the perceived expected returns are constant in time, but generally vary among assets with riskier assets commanding higher premia. On the other hand, under the objective measure $\mathbb{P}$ the return is given by

$$\mathbb{E}_t[r_{t+1}^{ex}(i)] = \underbrace{constant(i)}_{\text{Perceived risk premium}} + \underbrace{\mathbb{E}_t[r_{t+1}^{ex}(i)] - \mathbb{E}_t^S [r_{t+1}^{ex}(i)]}_{\text{Effect of time-varying mispricing}}. \tag{5}$$

The model generates such a perceived constant risk premium so that for example the term structure of interest rates is upward sloping on average. However, this paper does not offer a deep theory for such risk premia. Rather the modeling of this premium is reduced form and all of the novel and interesting effects come from expectational errors.

The structure of this paper is the following. I start by introducing the model and deriving some theoretical results for bond and currency returns. I then formulate empirical predictions, most of which are easily testable using linear regressions. Finally, I move to the empirical part and find that the model is able to generate coefficient values that are similar to those in the data.

2 A Term Structure Model with Expectational Errors

The model builds on Gourinchas and Tornell (2004) who only focus on currencies and matching some features of survey data. There are two symmetric countries, home and foreign, where the latter variables are denoted by stars. Moreover, let there be two (equivalent) probability measures P and S. Here P corresponds to objective probabilities as viewed by a rational econometrician. On the other hand, S represents the (homogeneous)[7] subjective beliefs of the agents. For simplicity I omit the P-symbol from expectations taken under rational beliefs.

Markets are complete. Under the subjective measure S, the home and foreign nominal SDFs follow symmetric (conditionally) log-normal processes[8]

$$log(M_{t,t+1}) \equiv m_{t,t+1} = -logR - \frac{\bar{\sigma}_{\epsilon}^2 \bar{\varphi}_t^2}{2} - \bar{z}_t - \frac{\sigma_{\epsilon}^2 \varphi_t^2}{2} - z_t - \bar{\varphi}_t \bar{\epsilon}_{t+1} - \varphi_t \epsilon_{t+1} \quad (6)$$

$$log(M^*_{t,t+1}) \equiv m^*_{t,t+1} = -logR - \frac{\bar{\sigma}_{\epsilon}^2 \bar{\varphi}_t^{*2}}{2} - \bar{z}_t - \frac{\sigma_{\epsilon}^2 \varphi_t^{*2}}{2} - z^*_t - \bar{\varphi}^*_t \bar{\epsilon}_{t+1} - \varphi^*_t \epsilon^*_{t+1}. \quad (7)$$

The shocks $\boldsymbol{\epsilon}_t = (\epsilon_t, \epsilon^*_t, \bar{\epsilon}_t)$ are independent and follow a (joint) normal distribution with mean zero and variances σ_{ϵ}^2, σ_{ϵ}^2 and $\bar{\sigma}_{\epsilon}^2$. z_t and z^*_t are country specific states and $\bar{z}_t$ is a state shared by both countries.

[7] Alternatively one can view S as the agents' average belief.

[8] The implications of such affine models for currencies have been studied by, for example, Backus et al. (2001), Lustig et al. (2011) and Lustig et al. (2017). However, these papers do not study the effects of expectational errors. Note that the assumption of symmetric pricing kernels implies that the model cannot explain persistent country level differences in returns. I leave extending the results to asymmetric pricing kernels to future work. Note that I use the term carry trade in a loose sense to refer to a strategy that borrows in a low interest rate currency and invests the proceeds in a high interest rate currency. However, the trading opportunity implied by the model is based on exploiting time-series rather than persistent cross-country violations of the uncovered interest parity. For a careful analysis of the different types of violations of uncovered interest parity and the related trading strategies, see Hassan and Mano (2017).

Under the objective measure, the states $\mathbf{z}_t = [z_t, z_t^*, \bar{z}_t]'$ follow the process

$$\mathbf{z}_t = \Lambda \mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \tag{8}$$

where

$$\Lambda = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \bar{\lambda} \end{bmatrix}. \tag{9}$$

Here $0 < \lambda < 1$ and $0 < \bar{\lambda} < 1$. On the other hand, the investors believe that these follow (i.e. their S-dynamics are given by)[9]

$$\mathbf{z}_t = \mathbf{l}_t + \boldsymbol{v}_t, \tag{10}$$

$$\mathbf{l}_t = \Lambda \mathbf{l}_{t-1} + \boldsymbol{\epsilon}_t. \tag{11}$$

Here $\boldsymbol{v}_t = [v_t, v_t^*, \bar{v}_t]'$ and $v_t, v_t^* \sim N(0, \sigma_v^2)$ and $\bar{v}_t \sim N(0, \bar{\sigma}_v^2)$, where each shock is independent. The investors erroneously believe that part of the shocks are transitory. This implies that the investors' expectations react to new information sluggishly.

Finally, the market prices of risk are given by

$$\Phi_t = [\varphi_t \ \bar{\varphi}_t \ \varphi_t^* \ \bar{\varphi}_t^*]' = \Phi_0 + \Phi_1 [z_t \ \ z_t^*]' \tag{12}$$

$$\Phi_0 = [\varphi_0 \ \ \bar{\varphi}_0 \ \ \varphi_0 \ \ \bar{\varphi}_0]' \tag{13}$$

$$\Phi_1 = \begin{bmatrix} \varphi_1 & 0 \\ \bar{\varphi}_1 & 0 \\ 0 & \varphi_1 \\ 0 & \bar{\varphi}_1 \end{bmatrix}. \tag{14}$$

Note that I for simplicity assume that only the local factor prices country specific and common shocks. However, in most of the empirical part I assume $\varphi_1 = \bar{\varphi}_1 = 0$ so that all of the return predictability will be due to expectational errors.

All of the key qualitative results of this paper could be derived by assuming the investors hold rational beliefs concerning the common shock. I

[9]Here the shocks are redefined.

could also allow for additional factors as well as make a separation between real and nominal pricing kernels by making an assumption on the inflation process (see e.g. Lustig et al. (2011) and Lustig et al. (2014)).[10]

The model yields a simple formula for the interest rate differential given by the following lemma

Lemma 1. *The log-interest rate differential is given by* $x_t \equiv i_t^* - i_t = z_t^* - z_t$. *The true law of motion for* x_t *is*

$$x_t = \lambda x_{t-1} + \epsilon_t^* - \epsilon_t \equiv \lambda x_{t-1} + \tilde{\epsilon}_t. \tag{15}$$

The perceived law of motion for x_t *is*

$$x_t = \tilde{l}_t + \tilde{v}_t, \tag{16}$$

$$\tilde{l}_t = \lambda \tilde{l}_{t-1} + \tilde{\epsilon}_t, \tag{17}$$

where $\tilde{v}_t = v_t^* - v_t$.

Proof. Note that

$$x_t = log(\mathbb{E}_t^S[exp(m_{t,t+1})]) - log(\mathbb{E}_t^S[exp(m_{t,t+1}^*)]). \tag{18}$$

All the states related to $m_{t,t+1}$ and $m_{t,t+1}^*$ are known. Using the mean of an exponential of a normal random variable

$$x_t = z_t^* - z_t. \tag{19}$$

The perceived and actual law of motion follow directly from the corresponding processes for z_t^* and z_t.

□

One implication of the lemma is that the λ coefficient corresponds to that in Gourinchas and Tornell (2004). This also implies that the states in the model can either represent deep economic factors or a reduced form characterization of a short rate process.

What is the economic interpretation of the states $\mathbf{z}_t$? One interpretation is that they determine the consumption growth rates in the two economies.

[10] Alternatively one could formulate the theoretical predictions for real pricing kernels and use data on real interest rates and exchange rates for the empirical part. For the potential effect of inflation risk on carry trade returns see Jylhä and Suominen (2010).

To see this, consider the following simple example. Assume the preferences in the two economies are given by

$$\sum_{t=0}^{\infty} \beta^t u(C_t) \tag{20}$$

$$\sum_{t=0}^{\infty} \beta^t u(C_t^*), \tag{21}$$

where

$$u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \tag{22}$$

and the utility functions are evaluated using the subjective measure S. Abstracting away from inflation, the log-SDF is given by

$$log(M_{t,t+1}) = log(\beta) - \gamma(c_{t+1} - c_t). \tag{23}$$

Assume for simplicity that the market prices of risk are constant. Then let

$$log(\beta) - \gamma(c_{t+1} - c_t) = -logR - \frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_0^2}{2} - \bar{z}_t - \frac{\sigma_\epsilon^2 \varphi_0^2}{2} - z_t - \bar{\varphi}_0 \bar{\epsilon}_{t+1} - \varphi_0 \epsilon_{t+1}. \tag{24}$$

That is $\bar{z}_t$ and z_t determine the expected growth rate of consumption. $\bar{z}_t$ is the shared part of the growth rate and z_t is the part specific to the home country. The process for $c_{t+1}^* - c_t^*$ is determined symmetrically. Now the SDF coincides with the one determined before. However, the data is unlikely to support such a direct interpretation of the states as the growth rate of the economy. Still, the literature has found some evidence that the factors in affine term structure models are driven by macroeconomic variables (see e.g. Ang and Piazzesi (2003)).

Note that I assumed the same shocks for the SDF and state variables. As in Gourinchas and Tornell (2004), I for simplicity assume the investors do not use the information in realized SDFs to update their beliefs about the economic state. One can view this correlation between the two shocks as a reduced form way of modeling an upward sloping yield curve. For the main model of this paper, it only affects the results through the constant terms of bond prices. These terms do not affect the main results of this paper that concern return predictability and yield volatilities.

Note that there is no clear agreement on the true cause of an upward sloping yield curve. The classic liquidity preference theory of Keynes (1936) loosely associates long-term securities with ”lower liquidity”. More specifically, an upward sloping curve can be obtained for example in models where consumption growth is negatively autocorrelated Campbell (1986). Moreover, it can arise in certain models with non-standard preferences such as in habit models (e.g. Wachter (2006)). Finally, it can be due to a negative correlation between consumption growth and inflation (e.g. Piazzesi and Schneider (2006)).[11]

The following gives a solution to the learning problem based on the standard recursion formulas for the Kalman filter (see e.g. Hamilton (1994)).

Proposition 1 (Learning Problem)**.** *Assume initial beliefs about $l_1, l_1^*, \bar{l}_1$ are normally distributed with l_1, l_1^* coming from the same distribution. Now the beliefs (are Gaussian and) evolve as*

$$\mathbb{E}_t^S[\boldsymbol{z}_{t+1}] = \begin{bmatrix} \lambda(1-k_t) & 0 & 0 \\ 0 & \lambda(1-k_t) & 0 \\ 0 & 0 & \bar{\lambda}(1-\bar{k}_t) \end{bmatrix} \mathbb{E}_{t-1}^S[\boldsymbol{z}_t] + \begin{bmatrix} \lambda k_t & 0 & 0 \\ 0 & \lambda k_t & 0 \\ 0 & 0 & \bar{\lambda}\bar{k}_t \end{bmatrix} \boldsymbol{z}_t, \tag{25}$$

$$\mathbb{E}_t^S[x_{t+1}] = \lambda(1-\tilde{k}_t)\mathbb{E}_{t-1}^S[x_t] + \lambda\tilde{k}_t x_t. \tag{26}$$

The formulas for k_t, $\bar{k}_t$, $\tilde{k}_t$ and the volatilities of the persistent components are given in the appendix. As $t \to \infty$, these estimators converge to steady-state values σ^2, $\bar{\sigma}^2$, $\tilde{\sigma}^2$, k and $\bar{k}$ given in the appendix.

The learning process for the foreign country is defined analogously. For the main results of this paper I for simplicity assume the estimators have converged to their steady-state values. Note that k_t and $\tilde{k}_t$ are generally different but converge to the same value.

To understand the key differences between subjective and objective expectations for the states, take the example of z_t. From the above proposition one has

$$\mathbb{E}_t^S[z_{t+1}] = (1-k_t)\lambda\mathbb{E}_{t-1}^s[z_t] + k_t\lambda z_t \tag{27}$$

[11] Also while nominal yield curves tend to be upward sloping, whether this is true for real yields is a more complicated question.

$$\mathbb{E}_t[z_{t+1}] = \lambda z_t. \tag{28}$$

If beliefs are rational $k_t = 1$ and the two expectations coincide. However, typically $0 < k_t < 1$ so that the subjective expectation is a weighted average of the last period expectation and the current value for the state. Effectively the biased measure underreacts to new interest rate shocks.

2.1 The Yield Curve

This section characterizes the generally time-varying yield curve as a function of the relevant state variables.

Proposition 2 (The yield curve)**.** *Denote the state variable* $\boldsymbol{Y}_t = [z_t, \bar{z}_t, \mathbb{E}_t^S[z_{t+1}], \mathbb{E}_t^S[\bar{z}_{t+1}]]'$. *The home logarithmic prices of zero coupon bonds are affine functions of* $\boldsymbol{Y}_t$ *and given by*

$$q_t(n) = A(n) + \boldsymbol{B}(n)'\boldsymbol{Y}_t, \tag{29}$$

where $A(n)$ *and* $\boldsymbol{B}(n)$ *are given by* $A(1) = -logR$ $\boldsymbol{B}(1) = [-1\ \ -1\ 0\ 0]'$ *and*

$$B_1(n) = -1 - \varphi_1(B_1(n-1) + k\lambda B_3(n-1))\sigma_\epsilon^2 - \bar{\varphi}_1(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))\bar{\sigma}_\epsilon^2, \tag{30}$$

$$B_2(n) = -1, \qquad B_3(n) = \lambda B_3(n-1) + B_1(n-1), \tag{31}$$

$$B_4(n) = \bar{\lambda}B_4(n-1) + B_2(n-1). \tag{32}$$

Finally $A(n)$ *is provided in the appendix.*

Assuming $\varphi_1 = \bar{\varphi}_1 = 0$ $\boldsymbol{Y}_t = [\hat{z}_t, \mathbb{E}_t^S[z_{t+1}], \mathbb{E}_t^S[\bar{z}_{t+1}]]$, *where* $\hat{z}_t = z_t + \bar{z}_t$ *and* $q_t(n) = A(n) + \boldsymbol{B}(n)'\boldsymbol{Y}_t$. *Here* $A(n)$ *and* $\boldsymbol{B}(n)$ *are given by* $A(1) = -logR$ $\boldsymbol{B}(1) = [-1\ 0\ 0]'$ *and*

$$B_1(n) = -1 \quad B_2(n) = \lambda B_2(n-1) - 1 \quad B_3(n) = \bar{\lambda}B_3(n-1) - 1. \tag{33}$$

Finally $A(n)$ *is given in the appendix.*

The interest rates are generally high when the factors and their expectations are high. Note that in a rational model the only factors determining the home yield curve would be z_t and $\bar{z}_t$.

2.2 Term Structure of Expected Carry Trade Returns

I now characterize the expected returns of the two currencies. The proposition derives results for both the relative returns on short-term bills in the two currencies as well as those for longer maturity bonds. As in Gourinchas and Tornell (2004) I for simplicity assume the agents hold correct views concerning the long-run component of the exchange rate.[12]

Proposition 3 (Term Structure of Expected Carry Trade Returns)**.** *Let the home relative (objective) expected return of foreign currency short-term bills be* $\Theta_t \equiv x_t + \mathbb{E}_t[s_{t+1}] - s_t$*. Further let the home relative (objective) expected return of long-term foreign currency bonds of maturity* $n \geq 2$ *be* $\Theta_t^n \equiv \mathbb{E}_t[q_{t+1}^*(n-1) - q_{t+1}(n-1)] - [q_t^*(n) - q_t(n)] + \mathbb{E}_t[s_{t+1}] - s_t$*. Now*

$$\Theta_t = \left[1 + \frac{\lambda k}{1-\lambda}\right]\left[\mathbb{E}_t x_{t+1} - \mathbb{E}_t^S x_{t+1}\right] + \mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} \Gamma_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{\infty} \Gamma_{t+j}\right], \tag{34}$$

where $0 < k < 1$ *and* $0 < \lambda < 1$ *and the perceived risk premium is given by*

$$\Gamma_t = -\frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^2}{2} - \frac{\sigma_\epsilon^2 \varphi_t^2}{2} + \frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^{*2}}{2} + \frac{\sigma_\epsilon^2 \varphi_t^{*2}}{2}. \tag{35}$$

The general expression for Θ_t^n *is given in the appendix. Assuming* $\varphi_1 = \bar{\varphi}_1 = 0$,

$$\Theta_t = \left[1 + \frac{\lambda k}{1-\lambda}\right]\left[\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}]\right] \tag{36}$$

$$\Theta_t^n = \frac{k\lambda^n}{1-\lambda}\left[\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}]\right]. \tag{37}$$

Now as $n \to \infty$, $\Theta_t^n \to 0$*. The term structure of expected carry trade returns is downward sloping. The long bond parity condition holds in the limit as the maturity of the bonds increase.*

[12] If I relaxed this assumption, the long-bond return parity condition would generally not hold exactly, but the term structure of carry trade returns would still be downward sloping. The determination of this permanent component is beyond the scope of this paper. For example in a model with a stationary real exchange rate, the permanent component of the exchange rate would be the permanent component of the log price differential in the two countries.

Proof: see appendix.

The general expression for the currency risk premium under the objective measure can be decomposed as follows

$$\underbrace{\Theta_t}_{\text{Currency premium}} = \underbrace{-\Gamma_t}_{\text{Risk premium differential}} \underbrace{+\mathbb{E}_t\left[\mathbb{E}^S_{t+1}\sum_{j=0}^{\infty}x_{t+1+j} - \mathbb{E}^S_t\sum_{j=0}^{\infty}x_{t+1+j}\right]}_{\text{Interest rate misperception effect}} + \tag{38}$$

$$\underbrace{+\mathbb{E}_t\left[\mathbb{E}^S_{t+1}\sum_{j=0}^{\infty}\Gamma_{t+1+j} - \mathbb{E}^S_t\sum_{j=0}^{\infty}\Gamma_{t+1+j}\right]}_{\text{Risk premium misperception effect}}. \tag{39}$$

(40)

A similar decomposition can be obtained for Θ^n_t. Rational models only include the risk premium channel. However, here misperceptions about economic states affect the actual currency premium in two ways. First they create time-variation in (actual) expected returns due to misperceptions about future interest rates. Second, they create additional time-variation in (actual) expected returns due to expectational errors about future risk adjustments.

For the rest of this paper I for simplicity assume $\varphi_1 = \bar{\varphi}_1 = 0$, which means that the perceived market prices of risk are constant. This means that time-variation in expected excess returns is only driven by misperceptions about future interest rates.

Looking at the simplified expressions given in the proposition, one can now see that the expected relative return on foreign short-term securities is positive when $\mathbb{E}_t[x_{t+1}] - \mathbb{E}^S_t[x_{t+1}] > 0$. This tends to happen when the interest rate differential x_t is high. Effectively the high interest rate currency is undervalued because the investors do not expect the high interest rate environment to persist.

The relative returns on foreign long maturity bonds are also positive when $\mathbb{E}_t[x_{t+1}] - \mathbb{E}^S_t[x_{t+1}] > 0$. However, they are decreasing in maturity n. Moreover, the LBP condition holds exactly in the limit as $n \to \infty$. The long-term bonds of the high interest currency are overvalued because high future short rates imply low returns for long-term bonds.

Note that when $k = 1$, that is assuming rational expectations, there is no time-variation in expected currency returns. As I explained before, in the simplified model all of the variation in risk premia are driven by expectational errors.

Finally, due to risk adjustments the general model does not typically satisfy the long-bond parity condition $\Theta_t^{\infty} = 0$. This is consistent with Lustig et al. (2017) who note that this condition does not arise naturally in many risk-based models. However, the simplified model naturally satisfies this condition.

3 Empirical Predictions

In this section I derive some empirical predictions of the model, most of which are easily testable using linear regressions.[13] I focus on bond risk premia and the term structure of expected carry trade returns. Additional implications are derived for yield volatilities and survey responses concerning currencies and long-term interest rates.

As discussed before I set the time-varying parts of market prices of risk to zero: $\varphi_1 = \bar{\varphi}_1 = 0$. This implies that all of the time-variation in excess returns will be due to expectational errors. It turns out that such a simple model can go a surprisingly long way in explaining bond and currency returns. All of the predictability results hinge only on the actual and perceived processes for the state variables, not on the other terms in the SDFs.[14]

3.1 Time-Variation in Term Premium

I start by considering term premium predictability. I focus on the following regression

$$q_{t+m}(n-m) - q_t(n) + q_t(m) = \alpha_1 + \beta_1(i_t(n) - i_t) + e_{t+m}, \tag{41}$$

where e_{t+m} is a standard zero mean error term. That is I predict excess bond returns using the slope of the yield curve. For simplicity assume the forecasters hold correct beliefs about the common shock.[15]Now

[13] All of the empirical predictions are naturally under the objective measure.

[14] Moreover, one benefit of considering regression slope coefficients instead of plain covariances is that the slope coefficients do not depend on factor variances.

[15] Results are similar assuming the misperceptions concerning this shock are similar than for the country specific shock.

$$Cov(q_{t+m}(n-m) - q_t(n) + q_t(m), i_t(n) - i_t) = \tag{42}$$
$$Cov(\sum_{i=0}^{n-m-1} \mathbb{E}^S_{t+m}[-z_{t+m+i}] + \sum_{i=0}^{n-1} \mathbb{E}^S_t[z_{t+i}] - \sum_{i=0}^{m-1} \mathbb{E}^S_t[z_{t+i}], m\sum_{i=0}^{n-1} \mathbb{E}^S_t[\frac{z_{t+i}}{n}] - mz_t).$$

I focus on one quarter predictability $m = 3$. Figure 1 shows the β_1 coefficient as a function of k for ten year bonds, $n = 120$.[16] When $k = 1$ expectations are rational and there is no predictability. On the other hand, when $k < 1$, a high yield curve slope predicts positive returns for long-term bonds. The effect is generally stronger for lower k, i.e. higher biases, though it starts to decline when biases become very large. Moreover, predictability is higher for larger values of the persistence parameter λ. Gourinchas and Tornell (2004) find support for values of k between 0.27 and 0.68 and values of λ between 0.95 and 1.01.

Figure 2 shows the effect for one year bonds, $n = 12$. One can see that the predictability coefficient is of similar magnitude as that for ten year bonds. However, the predictability coefficient starts declining already for higher values of k. I leave the consideration of more complicated predictability factors such as those employed by Cochrane and Piazzesi (2005) to future work.

3.2 Predictability in the Term Structure of Carry Trade Returns

I next consider the predictions for the term structure of expected carry trade returns. I consider predicting returns of the carry trade by the short-term interest differential.

Proposition 4 (Term Structure of Carry Trade Returns and Interest Rate Differential)**.** *Let* $\tilde{\Theta}_{t+1} \equiv i^*_{t+1} - i_{t+1} + s_{t+1} - s_t$ *and* $\tilde{\Theta}_{t+1}(n) \equiv q^*_{t+1}(n-1) - q_{t+1}(n-1) - (q^*_t(n) - q_t(n)) + s_{t+1} - s_t$ *. Consider the following regressions*

$$\tilde{\Theta}_{t+1} = \delta_0 + \delta_1 x_t + \varepsilon_{t+1} \tag{43}$$

$$\tilde{\Theta}^n_{t+1} = \delta^n_0 + \delta^n_1 x_t + \varepsilon^n_{t+1} \tag{44}$$

δ_1, δ^n_1 *are positive and*

[16] Results are unaffected by the variance of the shocks.

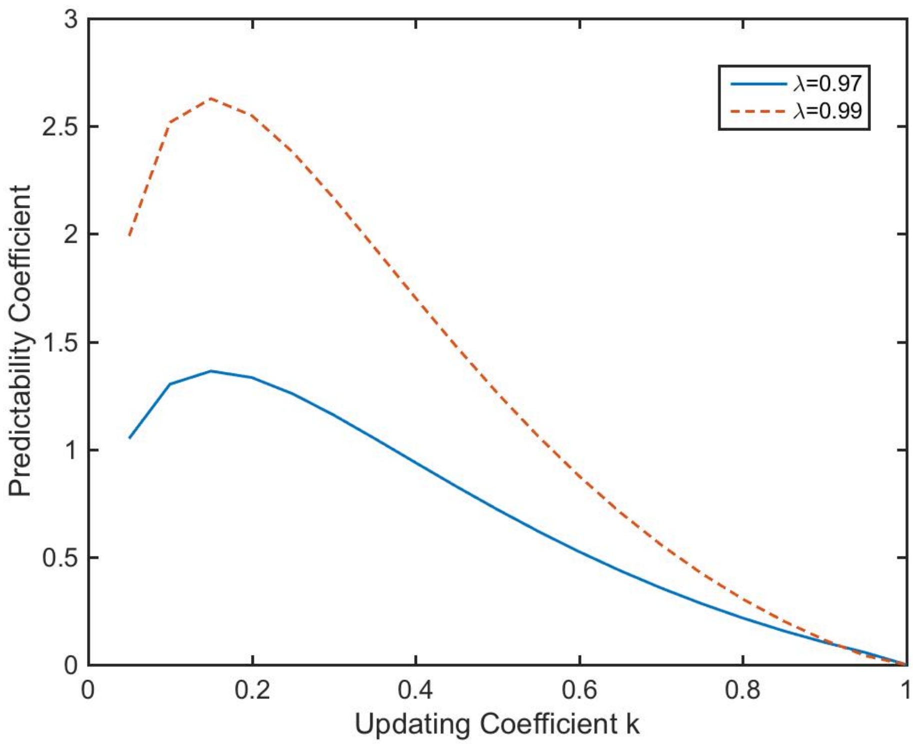

Figure 1 shows the slope coefficient of a regression of excess 10 year bond returns on yield curve slope (10 year minus 3 month bill) as a function of the updating coefficient k.

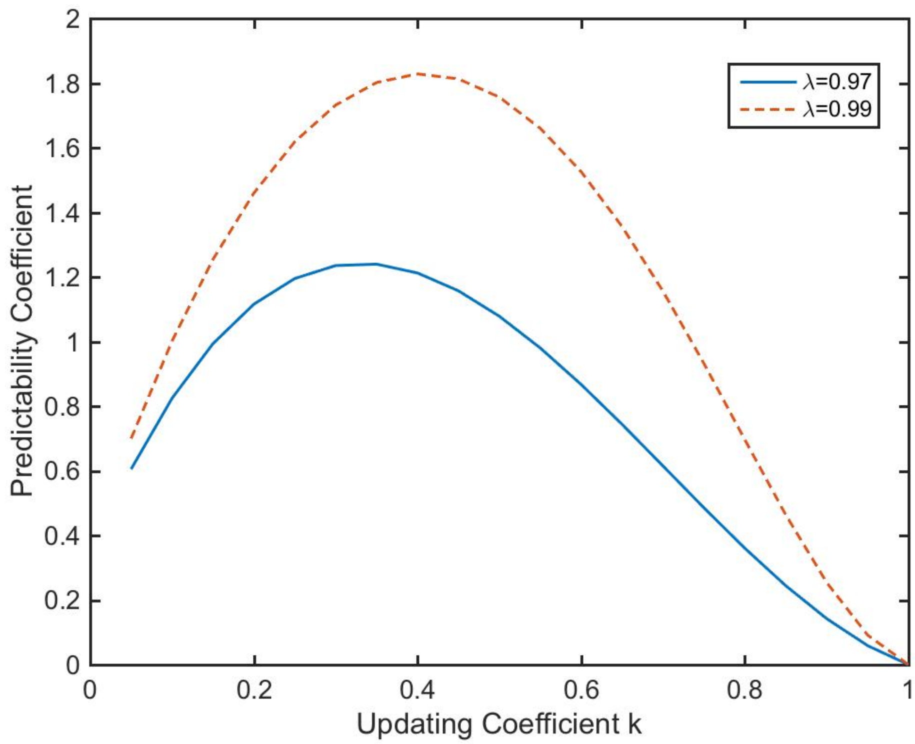

Figure 2 shows the slope coefficient of a regression of excess 1 year bond returns on yield curve slope (1 year minus 3 month bill) as a function of the updating coefficient k.

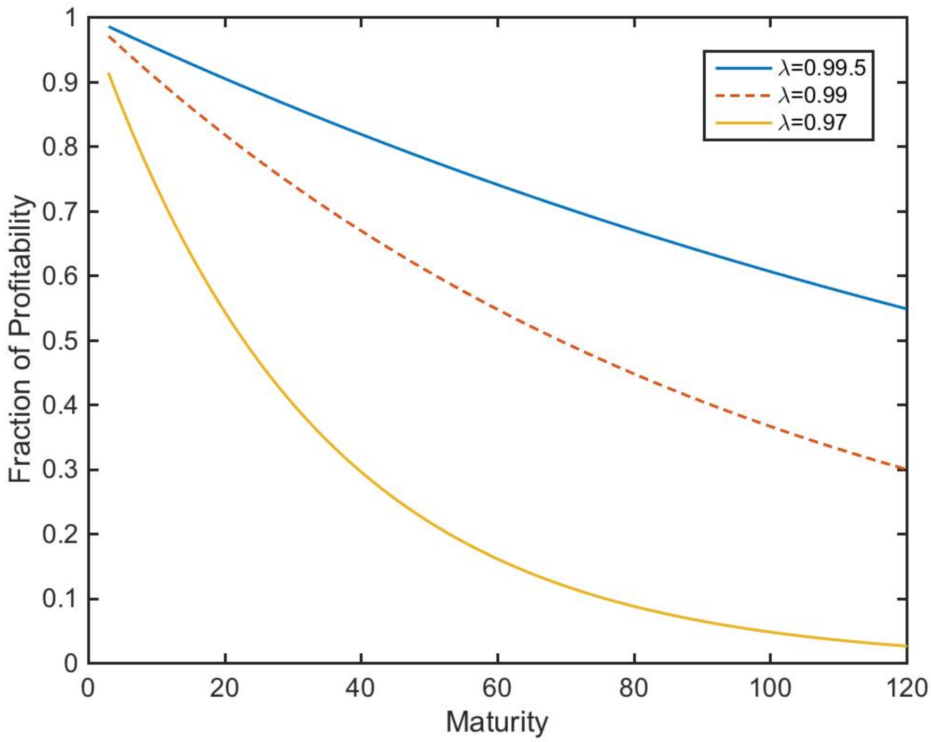

Figure 3 The figure shows the decay pattern of carry trade profitability for different maturity bonds.

$$\delta_1^n = \frac{k\lambda^n}{1-\lambda}\frac{(1-k)(\lambda-\lambda^3)}{1-(1-k)\lambda^2}. \tag{45}$$

δ_1^n decays at rate λ^n and approaches zero as $n \to \infty$.

Proof: see appendix.

A positive interest rate differential predicts positive carry trade returns for any maturity bonds. However, the effect is declining in the bond maturity n and there is no predictability in the limit $n \to \infty$. Figure 3 shows the relative slope coefficient as a function maturity n. The predictability of carry trade returns declines faster for smaller values of λ.

3.3 Predictions for Expectational Errors in Surveys

The process for short-term interest rate expectations is based on Gourinchas and Tornell (2004). Such a process has important implications for expectational errors concerning other instruments. In this section I study these implications for survey data on FX rates and long-term interest rates.

I start by considering predictions for survey data on FX rates. The following proposition shows that the model implies that forecasters underestimate the relative future strength of high interest rate currencies.

Proposition 5 (Matching Survey Data on Currencies)**.** *Consider the following regression*

$$s_{t+j} - \mathbb{E}_t^S[s_{t+j}] = \kappa_{10} + \kappa_{11} x_t + e_{t+j}. \tag{46}$$

Now $\kappa_{11} > 0$.

Proof: see appendix.

This prediction emerges also in Gourinchas and Tornell (2004), though here I show that it carries over to long horizons as well. Bacchetta et al. (2009) run regressions of this form and find strong support for a positive coefficient.[17] I replicate their regressions using a shorter more recent time series from Reuters and find similar yet weaker results.

I now move to consider long-term interest rates. Specifically, consider the regression

$$-(i_{t+m}(n) - \mathbb{E}_t^S[i_{t+m}(n)]) = \kappa_{20} + \kappa_{21}(i_t(n) - i_t) + e_{t+m}. \tag{47}$$

Assume the forecasters hold correct beliefs about the common shock. Now

$$Cov(-i_{t+m}(n) + \mathbb{E}_t^S[i_{t+m}(n)], i_t(n) - i_t) = \tag{48}$$
$$Cov(-\frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}_{t+m}^S[z_{t+m+i}] + \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}_t^S[z_{t+m+i}], \sum_{i=0}^{n-1}\mathbb{E}_t^S[\frac{z_{t+i}}{n}] - z_t).$$

This covariance tends to be positive and therefore the coefficient positive. This means that forecasters overpredict the importance of an upward sloping yield curve on future interest rates.

[17] Their coefficients are negative as they define the interest rate differential as home minus foreign rate.

For example following the results of Gourinchas and Tornell (2004) set $\lambda = 0.97$ and $k = 0.5$. Further consider one quarter predictability $m = 3$. Now $\kappa_{21} \approx 0.02$. This is smaller though of similar magnitude as the coefficients reported by Bacchetta et al. (2009).

3.4 Average Yield Curve

The average yield curve can be upward or downward sloping and is overall quite sensitive to parameter values. One typically needs $\varphi_0 < 0$ and $\bar{\varphi}_0 < 0$ to make the yield curve upward sloping on average. Figure 4 shows the average yield curve for a reasonable calibration of parameter values including a moderate misestimation of factor persistence.[18]

However, as argued before the modeling of the risk premia affecting the average yield curve is somewhat reduced form. The yield curve is upward sloping on average as longer term bonds are more sensitive to the same shocks that affect the stochastic discount factor (or marginal utility). However, this paper does not offer a deep theory for the nature of such shocks, mainly because the literature does not agree on the microfoundations behind an upward sloping yield curve.

3.5 Interest Rate Volatilities

The interest rate volatility formulas are given in the appendix. Note that for short-term rates

$$\mathbb{V}ar(12i_t) = 12^2\mathbb{V}ar(z_t) + 12^2\mathbb{V}ar(\bar{z}_t). \tag{49}$$

In the empirical section one will see that in the data short-rate volatility is roughly 3% and 10 year rate volatility approximately 2%.

To obtain a monthly volatility of 3% one can for example set $Var(z_t) = Var(\bar{z}_t) = 0.0017^2$.[19] Then figure 5 below shows the volatility of the 10 year yield as a function of the updating coefficients k and $\bar{k}$ when $\lambda = \bar{\lambda} = 0.97$. One can see that biased beliefs increase the volatility of long-term interest rates, but the effect is weak for small biases (high k and $\bar{k}$). Long-term yields are less volatile than short-term yields. When beliefs are fully rational ($k = 1$, $\bar{k} = 1$), 10 year yield volatility is roughly 1%. On the other hand, when

[18] $\lambda = 0.97$, $\bar{\lambda} = 0.97$, $k = 0.62$, $\bar{k} = 1$, $\varphi_0 = -4.93$, $\bar{\varphi}_0 = -4.93$, $\sigma_\epsilon = 0.006$, $\sigma_{\bar{\epsilon}} = 0.006$, $\eta = 0.5$, $\bar{\eta} = 0.5$, $log(R) = 0.003$.

[19] This simple calibration implies a cross-country short rate correlation of 0.5. This is close to the average value for the rest of the G10 countries against the US.

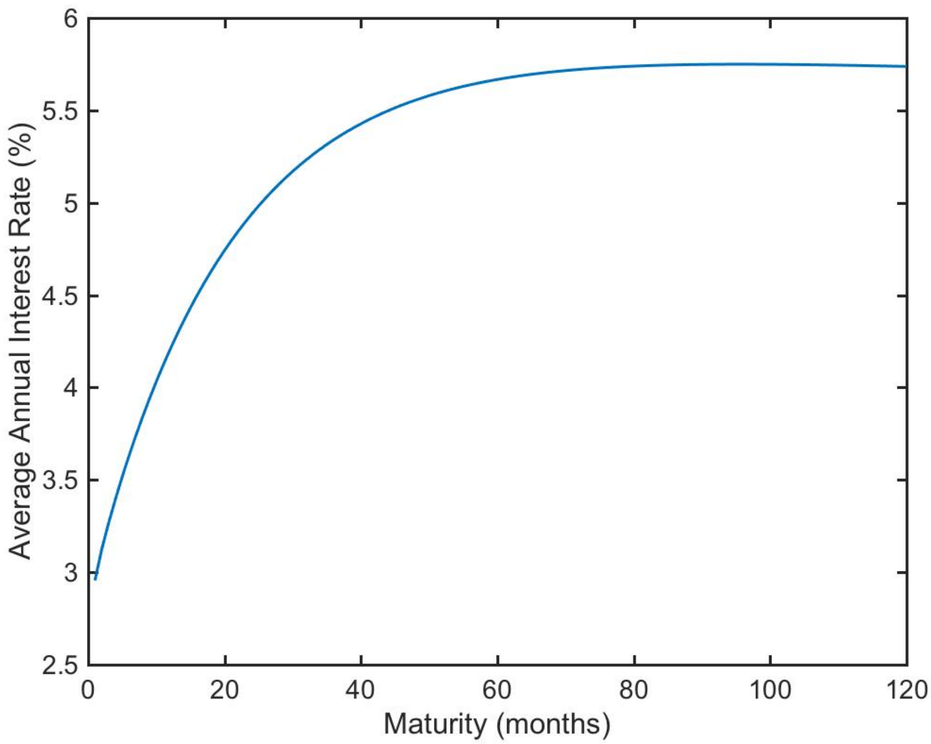

Figure 4 shows the average yield curve for one calibration of parameter values.

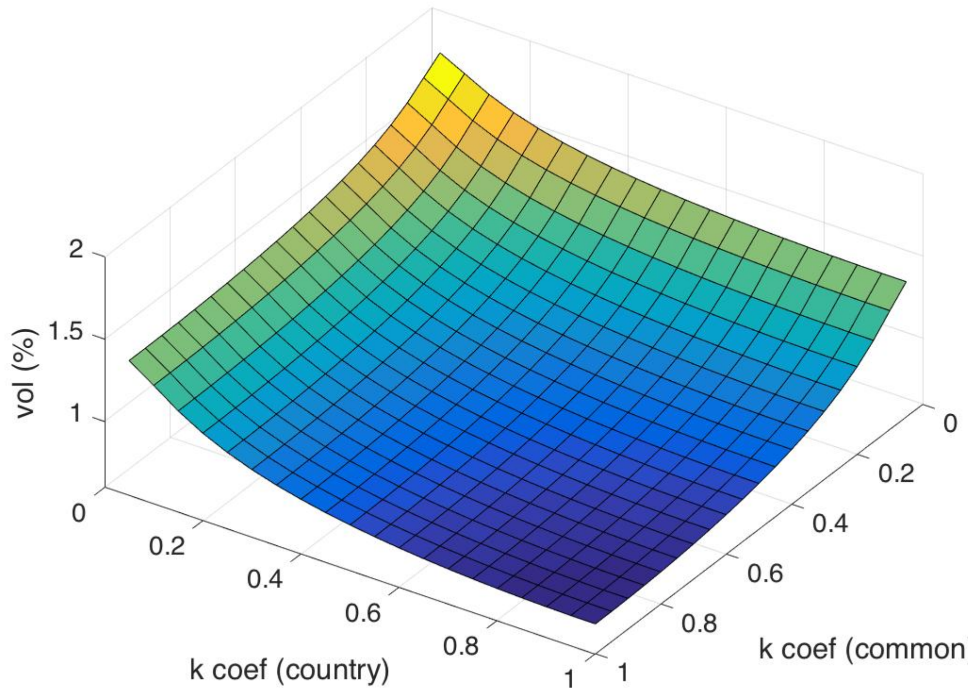

Figure 5 Monthly volatility of ten year interest rates as a function of the updating coefficients k (country specific) and $\bar{k}$ (common), $\lambda = \bar{\lambda} = 0.97$

beliefs are fully irrational ($k = 0$, $\bar{k} = 0$), 10 year yield volatility is roughly 2%.

Figure 6 shows the same long-term rate volatility when $\lambda = \bar{\lambda} = 0.99$. One can see that reaching a volatility of 2% requires smaller, more realistic, biases when the factors are more persistent. Moreover, when biases are large long-term rates can even be more volatile than short-term rates.

A well-known puzzle is that long-term yields are more volatile than standard representative agent models predict (e.g. Shiller (1979)). Similarly assuming rational expectations $k = \bar{k} = 1$, the model cannot match the observable levels of long-term yield volatility. However, as one will also see in the empirical part, bias levels close to those calibrated by Gourinchas and Tornell (2004) imply realistic volatility levels for both short and long yields. Therefore expectational errors may provide at least a partial resolution to this yield volatility puzzle.

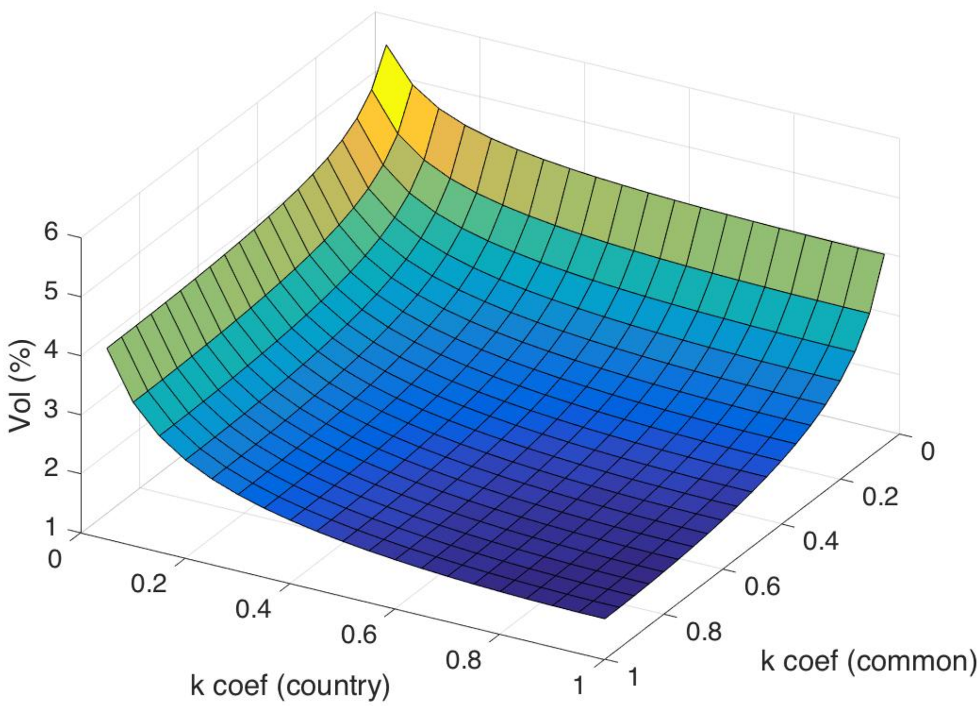

Figure 6 Volatility of ten year interest rates as a function of the updating coefficients k (country specific) and $\bar{k}$ (common), $\lambda = \bar{\lambda} = 0.99$.

3.6 Delayed Overshooting, Volatility and Persistence

The main model implies an identical exchange rate process to that in Gourinchas and Tornell (2004). They show that this process can account for these three puzzles. Delayed overshooting refers to the fact that the response of the FX rate to interest rate shocks is hump-shaped (Eichenbaum and Evans, 1995). The volatility puzzle corresponds to the fact that the exchange rate is much more volatile than predictable currency returns or interest rate differentials. The persistence puzzle refers to the fact that interest rate differentials are much more persistent than exchange rate changes.

4 Empirical Evidence

Most of the qualitative predictions of the model have been confirmed in the previous literature. Therefore I focus on checking whether the model can quantitatively match the data. I find that for bias levels similar to those calibrated by Gourinchas and Tornell (2004), the model implies realistic coefficient values related to time-variation in the term premium and term structure of carry trade returns. On the other hand, my results concerning expectational errors in FX forecasts are positive yet weaker than those obtained by the previous liteture. However, in comparison to for example Bacchetta et al. (2009) these results are based on a shorter more recent time-series of FX forecasts.

I focus on the G10 currencies of Australia, Canada, Germany, Japan, Norway, New Zealand, Sweden, Switzerland, U.K. and U.S. I use Datastream and FRED to obtain data on end of month FX rates and rates on 3 month government securities. I use two data sources for longer term bonds. First, I use the detailed dataset on zero coupon international bond yields constructed by Wright (2011). This dataset covers zero coupon yields for bonds with 40 different maturities between 3 months and 10 years. Secondly, I use Citigroup local currency 10 year bond indices available for all countries except New Zealand and Norway.[20]

The downside of using bond indices is that they are based on coupon bonds, while the theoretical predictions are for zero coupon bonds. Moreover, the indices are available only for certain maturity bonds; here I focus on 10 year bonds. On the other hand, the zero coupon bond yield curves are obtained using certain interpolation procedures (see Wright (2011)).

[20] These bond indices have been used as a robustness check for example by Dahlquist and Hasseltoft (2013).

	AUS	CAN	GER	JAP	NOR	NZ	SWE	CH	UK	US
					ZC Bonds					
Start	87	86	77	85	98	90	92	88	79	77
End	09	09	09	09	09	09	09	09	09	09
Obs	268	281	389	293	137	233	198	257	365	389
					Indices					
Start	84	84	94	87	NA	NA	92	84	84	84
End	09	09	09	09	NA	NA	09	09	09	09
Obs	301	301	192	268	NA	NA	203	301	301	301

Table 1 Start and end dates for data on zero coupon bonds and bond indices.

Therefore the advantage of the indices is that the subsequent returns are free from such approximations. Table 1 shows the time periods for the data.[21]

4.1 Term Premium Predictability

I first consider term premium predictability using the regression

$$q_{t+m}(n-m) - q_t(n) + q_t(m) = \alpha_1 + \beta_1(i_t(n) - i_t) + e_{t+m}. \tag{50}$$

As discussed before the model predicts a positive slope coefficient $\beta_1 > 0$. Following the theoretical section, I focus on quarterly predictability $m = 3$. As in the theoretical calibrations, the yield spread is measured on a quarterly basis (that is annual levels are divided by four).

Table 2 shows the results for ten year bonds using zero coupon bonds. The slope coefficient is positive for all countries except Norway. A high slope of the yield curve predicts high excess returns for long-term bonds. Most of the coefficients are above one. This is consistent with Campbell and Shiller (1991) who find that when the yield spread is large, long-term yields tend to fall. Contrasting the empirical results to the theoretical predictions given by Figure 1, one can see that the model is able to create slope coefficients that are of similar magnitude to those in the data, at least when factor persistence is high.

The average slope coefficient across countries is 2.2. For $k = 0.4$[22] and $\lambda = 0.97$ the model predicts a slope coefficient of 1 and for $k = 0.4$ and $\lambda = 0.99$ a slope coefficient of 1.8.

[21] The zero coupon bond dataset of Wright (2011) ends in 2009. For comparability I exlude the most recent period also from the bond index data.

[22] This is the mean value in Gourinchas and Tornell (2004).

	$\alpha_1(3)$	$s.e$	$\beta_1(3)$	$s.e$	R^2
AUS	0.0081	0.0052	2.27	1.46	0.027
CAN	0.0039	0.0054	2.46*	1.34	0.037
GER	0.0016	0.0047	1.88**	0.84	0.029
JAP	-0.0032	0.0081	4.59**	2.06	0.063
NOR	0.0060	0.0064	-0.43	0.43	0.001
NZ	0.0070	0.0059	1.59	1.93	0.007
SWE	0.0120	0.0130	0.73	2.93	0.001
CH	0.0005	0.0043	2.54***	0.92	0.077
UK	0.0071	0.0048	1.56	1.18	0.012
US	-0.0110	0.0082	4.48***	1.48	0.066

Table 2 shows the results from regressing quarterly 10 year bond excess returns, based on zero coupon bonds, on yield curve slope (10 year yield minus 3 month yield). The standard errors are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

Table 3 shows the results for 1 year bonds. The slope coefficient is positive with the exception of Norway and Sweden. The average value for the slope coefficient is 0.8. For $k = 0.4$ and $\lambda = 0.97$ the model implies a slope coefficient of 1.2.

Table 4 depicts the results for the 10 year bond indices. It is seen that the slope coefficients are similar to those obtained using the zero coupon bonds. I end the section by concluding that overall the model seems to be able to generate term premium predictability that is roughly similar to that in the data.

4.2 Carry Trade Returns

I now move to consider predictability in the term structure of carry trade returns. I focus on predictability on a quarterly horizon.[23]For different maturity bonds n, I run the following regression

$$\tilde{\Theta}_{t+1}^n = \delta_0^n + \delta_1^n x_t + \varepsilon_{t+1}^n. \tag{51}$$

That is, I explain carry trade returns for different maturity bonds using the short-term interest rate differential. For short maturity bills the slope coefficient takes the same value as that in Gourinchas and Tornell (2004), who showed that typical values of k and λ can replicate the coefficient seen

[23]Results are similar for a monthly horizon (possible with bond indices).

	$\alpha_1(3)$	*s.e*	$\beta_1(3)$	*s.e*	R^2
AUS	0.00099*	0.00057	0.82**	0.386	0.061
CAN	0.00107*	0.00056	1.26***	0.461	0.074
GER	0.00054	0.00043	1.08***	0.179	0.155
JP	0.00060	0.00029	0.55	0.670	0.010
NO	0.00006	0.00084	-0.82	0.656	0.020
NZ	0.00073	0.00059	0.43	0.512	0.008
SWE	0.00191*	0.00060	-0.16	0.494	0.001
CH	0.00010	0.00048	1.25***	0.421	0.070
UK	0.00012	0.00051	0.97***	0.266	0.087
US	-0.00095	0.00088	2.26***	0.553	0.110

Table 3 shows the results from regressing quarterly 1 year bond excess returns based on zero coupon bonds on yield curve slope (1 year yield minus 3 month yield). The standard errors are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

	AUS	CAN	GER	JP	SWE	CH	UK	US
α_1	0.0055	0.0068	0.0262	-0.0250	0.0054	-0.0002	0.0047	-0.0018
s.e.	0.0039	0.0039	0.0249	0.0119	0.0114	0.0040	0.0037	0.0065
β_1	0.44	0.30	-1.62	8.18***	1.78	2.83***	1.96**	2.80**
s.e.	1.01	0.99	4.32	2.51	2.38	0.87	0.95	1.14
R^2	0.001	0.0004	0.002	0.108	0.008	0.108	0.037	0.035

Table 4 shows the results from regressing quarterly 10 year bond excess returns, based on the bond indices, on yield curve slope (10 year yield minus 3 month yield). The standard errors are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

	AUS	CAN	GER	JAP	SWE	CH	UK
$\delta_{10}(3)$	-0.0004	0.0024	0.0091	0.0297***	0.0126*	0.0050	0.0025
s.e.	0.0071	0.0041	0.0053	0.0085	0.0072	0.0068	0.0065
$\delta_{11}(3)$	1.21*	1.02**	0.67	3.63***	1.94	0.58	1.23
s.e.	0.66	0.52	1.09	1.07	1.28	1.28	1.26
R^2	0.021	0.015	0.004	0.068	0.005	0.028	0.010
$\delta_{10}(120)$	-0.0016	0.0060	0.0145	0.0237***	0.0052	0.0041	0.0009
s.e.	0.0086	0.0053	0.0105	0.0074	0.0062	0.0070	0.0084
$\delta_{11}(120)$	0.67	-0.47	0.63	3.59***	1.77*	-0.94	0.44
s.e.	1.03	0.73	2.36	1.17	1.01	1.43	1.51
R^2	0.004	0.002	0.001	0.058	0.023	0.010	0.001

Table 5 shows the results from regressing ”carry trade” returns implemented using either 3 month bills or long-term bonds (bond indices) on short-term interest differential (base currency USD). The standard errors are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

in the data. Therefore I focus on the decay pattern in the term structure of expected carry trade returns.

Table 5 shows the results using 3 month bills and 10 year bond indices. The average slope coefficient for currency returns using short-term bills is 1.5, while the average coefficient for curreny returns using bond indices is 0.8. This implies a λ coefficient of 0.995, which is rather high but still in the range of coefficient values obtained by Gourinchas and Tornell (2004).

Figure 5 shows the decline pattern for zero coupon bonds along with 95% confidence intervals. Here the currency returns and interest rates are averaged across different countries. Figure 6 shows the same pattern excluding the smallest countries New Zealand, Norway and Sweden. The decline in the predictability of carry trade returns is smaller for the zero coupon bond dataset implying a λ coefficient close to one. However, the slope coefficients for the longest maturity bonds are not statistically significant.

The above regressions do not constitute a direct trading strategy. Therefore now consider the following simple ”carry trade”. Assume one buys a foreign bond and sells a home bond when the short-term interest rate differential is positive: $x_t > 0$. Similarly assume one does the oppositive when the interest rate differential is negative: $x_t < 0$. Further consider two different versions of this strategy: one using short-term bills and one using 10 year bonds. It can be shown that the profits from this strategy should decline at rate λ^n

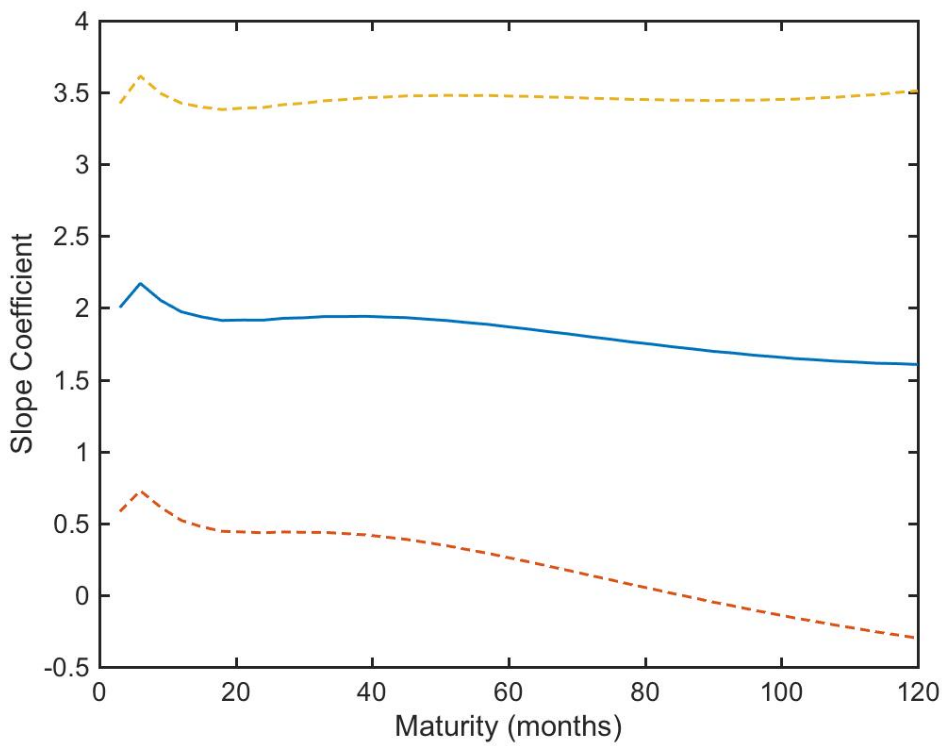

Figure 7 Term structure of expected carry trade returns. The figure shows the slope coefficient obtained by regressing the average relative return of foreign currency investments implemented with different maturity (zero coupon) bills/bonds on the average short-term interest rate differential (base currency USD). It also shows the 95% confidence bounds, based on robust standard errors, for the coefficient.

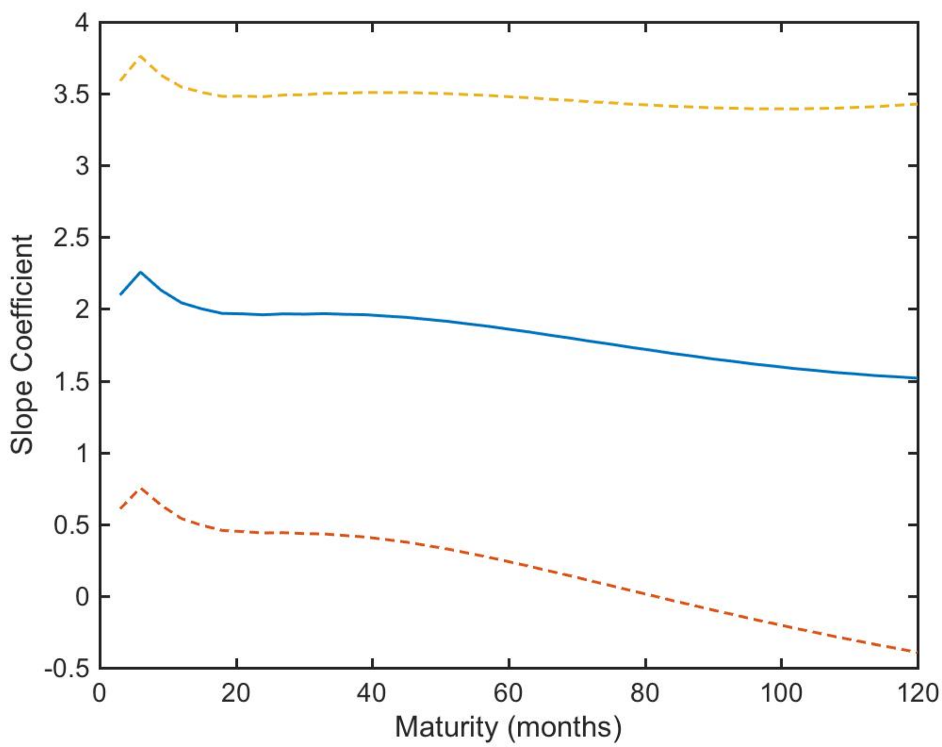

Figure 8 Term structure of expected carry trade returns excluding Norway, New Zealand and Sweden. The figure shows the slope coefficient obtained by regressing the average relative return of foreign currency investments implemented with different maturity (zero coupon) bills/bonds (base currency USD) on the average short-term interest rate differential. It also shows the 95% confidence bounds, based on robust standard errors, for the coefficient.

similarly to the regression coefficients.[24]

The average annualized log return for the short-term bill strategy is 3.31%. For the zero coupon bonds, the strategy using long-term bonds yields lower profits than the one using short-term bills for all countries except Switzerland. The average annualized return for the 10 year bond strategy is 2.68%. This implies a λ coefficient of 0.998.[25] On the other hand, for bond indices the mean annualized return from the long-term bond strategy is 1.76%, which implies a λ coefficient of 0.995.[26]

Using a longer sample period, a monthly holding period and panel regressions, Lustig et al. (2017) find a slope coefficient of 1.96 for short bills and 0.64 for 10 year bonds. This implies a λ coefficient of 0.991, which is closer to typical values of interest rate persistence observed in the data.

I conclude that the model is capable of replicating the decay pattern in expected carry trade returns seen in the data. However for typical values of λ, the model slightly overpredicts the decay pattern seen in the data. However, this might be partly related to the particular sample period used in this paper, as the results of Lustig et al. (2017) suggest a lower value for λ. Addional support for the long bond parity condition is given by Chinn and Meredith (2005) and Zviadadze (2016).

The model could also be modified to generate a milder decline in the term structure of carry trade, while allowing for a less persistent interest rate process. This could be obtained by assuming the short-term interest rate process is driven by a persistent and transitory component and the agents hold incorrect views about the persistent component. Furthermore the results depend on the assumption that the shocks follow AR(1) processes. I leave extending the derivations to more general shock processes to later work. Finally, the numerical results would change slightly if I relaxed the assumption that agents hold correct views concerning the long-run component of exchange rates.

[24] The expected profit from the strategy is
$\frac{k\lambda^n}{1-k}\mathbb{E}\Big[[\mathbb{E}_t[x_t]-\mathbb{E}_t^S[x_t]]I(x_t>0)-[\mathbb{E}_t[x_t]-\mathbb{E}_t^S[x_t]]I(x_t<0)\Big]$.

[25] The mean of the lambda coefficients calculated separately for each country is similar.

[26] The volatilities of the long bond strategies are only slightly higher than the volatilities of the short bill strategies, implying that convexity adjustments do not create large differences in the average gross returns of the two strategies. Overall, the distinction between log and gross returns is not important for this paper as long as one uses one of these coherently (here I formulate both the theoretical predictions as well as empirical results for log returns).

4.3 Survey Data on Currency Expectations

The key expectational predictions of the model are that forecasters (i) underweight the importance of recent interest rate shocks on future interest rates, (ii) underpredict the future strength of high interest rate currencies and (iii) overstate the importance of an upward sloping yield curve on future interest rates. These predictions have been verified by, for example, Froot and Frankel (1989), Gourinchas and Tornell (2004) and Bacchetta et al. (2009). Here I study the predictions for FX rates, given in proposition 5, using an alternative dataset.

More specifically, I use the FX polls data collected by Reuters. This data is based on a survey of over 150 strategists. I focus on the USD exchange rates of the major currencies Australian dollar, Canadian dollar, Euro, Japanese Yen, Swedish Krona, Swiss Franc and British Pound. The monthly panel is unbalanced. The longest data series for the Yen and the Pound range from 1995 to 2017 while the other series are clearly shorter and the data for the Krona starts only in 2012.

As in Gourinchas and Tornell (2004) and Bacchetta et al. (2009) I focus on the average forecasts of all institutions.[27] However, I use an equally weighted mean forecast rather than a consensus forecast obtained using different importance weights. Similar to Bacchetta et al. (2009) I apply the regression

$$s_{t+j} - \mathbb{E}_t^S[s_{t+j}] = \kappa_{10} + \kappa_{11} x_t + e_{t+j}. \tag{52}$$

That is, I explain expectational errors in FX forecasts using the short-term interest rate differential. As verified in proposition 5, the model predicts that $\kappa_{11} > 0$ [28], that is forecasters underpredict the future strength of high interest rate currencies.

This result has been confirmed in the previous literature and has received wide attention. For example Froot and Frankel (1989) argue that these biases can explain all of the violations of uncovered interest rate parity. While the early studies are based on fairly limited amounts of data, Bacchetta et al. (2009) find strong support for this prediction using long time-series.[29] Table

[27]The effect of disagreement on bond and currency risk premia has been considered in other papers. For bond markets see Buraschi and Whelan (2012), Ehling et al. (2013), Xiong and Yan (2010) and Giocoletti et al. (2016). For currency markets see Buraschi et al. (2010).

[28]However note that the model that assumes symmetric countries cannot be directly used to compute theoretical equivalents to empirical slope coefficients.

[29]Their coefficients are negative because they regress the exchange rate on home minus foreign interest rate.

6 shows the results using the alternative dataset from Reuters for horizons of 1,3,6 and 12 months.

The results are quite weak for short horizons, though the mean slope coefficients are positive for all horizons.[30] For the annual horizon the coefficients are positive with the exception of Canada. The average slope coefficient is 0.9.[31] These coefficients are somewhat smaller than those obtained by Bacchetta et al. (2009). Moreover, most of them are not statistically significant. This is in contrast to the results reported by Bacchetta et al. (2009) using longer time-series, who report a statistically significant positive coefficient for 6 out of 7 FX rates. Their coefficient values tend to be larger and significant also for short horizon predictability. Still the results point to the same direction as those obtained by Bacchetta et al. (2009).

A broader comparison between the results obtained using different datasets is beyond the scope of this paper. However, it would be interesting to see whether these results are specific to the recent data published by Reuters or whether FX forecasts have actually become more accurate over time. Some recent results suggest that these biases have not strongly diminished over time. For example Koijen et al. (2014) find that respondents on the World Economic Survey seem to expect negative returns on standard currency carry trade strategies.

Further note that for most part the survey data period coincides with a somewhat special period during which carry trade strategies have provided weak returns.[32]Finally, the use of survey data has well known caveats. For example, it is not clear whether these forecasts correspond to those of market participants.

4.4 Interest Rate Volatilities

Finally, I consider the model implications for volatilities of different maturity interest rates. Table 7 shows the volatility of 3 month and 10 year interest rates. The average monthly short rate volatility is roughly 3%. On the other hand, the average 10 year rate volatility is approximately 2%. The theoretical predictions for long-term rates are shown in Figures 5 and 6 given that the volatility of the short-term rate is calibrated to 3%.

[30]This holds also excluding the short times series for Sweden.

[31]The mean coefficient is 0.5 excluding Sweden.

[32]Moreover, some authors such as Engel (2016) voluntarily exclude the data period after the financial crises due to the possibility of changes in the driving processes of interest rates and currencies.

	AUS	CAN	EUR	JAP	SWE	CH	UK
$\kappa_{10}(1)$	-0.0091	0.0019	-0.0046**	0.0025	-0.0155	0.0098***	-0.0022
s.e.	0.0079	0.0027	0.0021	0.0031	0.0218	0.0036	0.0021
$\kappa_{11}(1)$	0.147	-0.109	-0.029	0.086	2.697	0.332*	-0.041
s.e.	0.317	0.289	0.181	0.097	1.755	0.176	0.154
R^2	0.002	0.001	0.0001	0.003	0.047	0.016	0.0003
$\kappa_{10}(3)$	-0.0276**	0.0060**	-0.0085	0.0103**	-0.0121	0.0214***	-0.0063**
s.e.	0.0122	0.0026	0.0028	0.0043	0.0270	0.0049	0.0030
$\kappa_{11}(3)$	0.558	-0.508*	0.076	0.205	2.151	0.827***	0.010
s.e.	0.537	0.264	0.247	0.135	1.956	0.236	0.209
R^2	0.022	0.020	0.0007	0.015	0.031	0.076	0.002
$\kappa_{10}(6)$	-0.0383***	0.0084***	-0.0119***	0.0240***	-0.0141	0.0375***	-0.0073*
s.e.	0.0128	0.0028	0.0040	0.0064	0.0303	0.0059	0.0045
$\kappa_{11}(6)$	0.661	-0.701***	-0.037	0.572	1.934	1.572***	0.012
s.e.	0.520	0.287	0.315	0.203	2.142	0.310	0.284
R^2	0.034	0.041	0.0001	0.077	0.025	0.194	0.00002
$\kappa_{10}(12)$	-0.0368***	0.0120***	-0.0156**	0.0401***	-0.0229	0.0530***	-0.0010
s.e.	0.0129	0.0037	0.0074	0.0129	0.0256	0.0094	0.0073
$\kappa_{11}(12)$	0.330	-1.491***	0.555	1.027***	2.100	1.999***	0.311
s.e.	0.393	0.437	0.552	0.332	2.825	0.677	0.410
R^2	0.010	0.158	0.023	0.134	0.029	0.197	0.010

Table 6 shows the results for predicting expectational errors in FX forecasts by the short-term interest differential for horizons of 1, 3, 6 and 12 months. The standard errors are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

	AUS	CAN	GER	JAP	NOR	NZ	SWE	CH	UK	US
3m yield vol	3.6	3.0	2.7	2.5	2.0	2.1	2.2	2.6	3.9	3.2
10y vol	2.8	2.2	1.7	2.0	0.8	1.6	2.1	1.3	3.1	2.6
Ratio	1.3	1.4	1.6	1.3	2.6	1.3	1.0	2.0	1.3	1.2

Table 7 shows the volatility of 10 year and 3 month interest rates (%) along with their ratio.

For example when $k = \bar{k} = 0.4$ and $\lambda = 0.97$ the model implies a 10 year rate volatility of 1.1%. On the other hand, when $k = \bar{k} = 0.4$ and $\lambda = 0.99$ the model implies a 10 year rate volatility of 2.4%. One can see that the model is able to generate realistic levels of volatility for both short- and long-term rates though this requires a relatively high value for the persistence parameter λ.

5 Conclusion

I provide an account of time-series predictability in bond and currency returns based on expectational errors alone. The model builds on Gourinchas and Tornell (2004) who find that forecasters tend to overweight the impact of transitory interest rate shocks and therefore underreact to interest rate innovations. They argue that such biases can explain the time-variation in expected currency returns. I argue that the same biases can also explain the time-variation in expected bond returns as well as the joint variation in bond and currency premia.

Importantly the biases work in opposite directions for bonds and currencies. The relative prices of currencies are increasing and the relative prices of long-term bonds decreasing in expected short rates. Therefore high interest rate currencies tend to be underpriced but the long-term bonds of these same currencies overpriced. This provides a novel explanation for the fact that the term structure of expected carry trade returns is downward sloping.

Traditional asset pricing models assume rational beliefs and explain *all* time-variation in expected returns as variation in required compensation for risk. On the other hand in the main model of this paper *all* time-variation in expected returns is due to expectational errors. Ultimately both types of models might seem too extreme. Future empirical research should attempt to carefully separate the relative quantitative importance of variation in risk premia from variation in expectational errors.

References

Ang, A. and Piazzesi, M. (2003). A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, 50:745–787.

Bacchetta, P., Mertens, E., and Wincoop, E. V. (2009). Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance*, 28:406–426.

Bacchetta, P. and Wincoop, E. V. (2010). Infrequent portfolio decisions: A solution to the forward discount puzzle. *American Economic Review*, 100:870–904.

Backus, D., Foresi, S., and Telmer, C. (2001). Affine models of currency pricing: Accounting for the forward premium anomaly. *Journal of Finance*, 56:279–304.

Backus, D., Gregory, A., and Telmer, C. (1993). Accounting for forward rates in markets for foreign currency. *Journal of Finance*, 48(5):1887–1908.

Bansal, R. and Shaliastovich, I. (2012). A long-run risks explanation of predictability puzzles in bond and currency markets. *Review of Financial Studies*, 26(1):1–33.

Buraschi, A., Beber, A., and Breedon, F. (2010). Difference in beliefs and currency risk premia. *Journal of Financial Economics*, 98:415–438.

Buraschi, A. and Whelan, P. (2012). Term structure models and differences in beliefs. Working Paper.

Burnside, C., Bing, H., Hirshleifer, D., and Yue Wang, T. (2011). Investor overconfidence and the forward premium puzzle. *Review of Economic Studies*, 78(2):523–558.

Campbell, J. (1986). Bond and stock returns in a simple exchange economy. *Quarterly Journal of Economics*, 101(3):785–804.

Campbell, J. and Shiller, R. (1991). Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies*, 58:495–514.

Chinn, M. and Meredith, G. (2005). Testing uncovered interest rate parity at short and long horizons during the post-bretton woods era. NBER Working Paper 11077.

Cieslak, A. (2016). Short-rate expectations and unexpected returns in treasury bonds. Working Paper.

Cochrane, J. and Piazzesi, M. (2005). Bond risk premia. *American Economic Review*, 95(1):138–160.

Dahlquist, M. and Hasseltoft, H. (2013). International bond risk premia. *Journal of International Economics*, 90:17–32.

Ehling, P., Gallmeyer, M., Heyerdahl-Larsen, C., and lleditsch, P. (2013). Disagreement about inflation and the yield curve. Working Paper.

Eichenbaum, M. and Evans, C. (1995). Some empirical evidence on the effects of monetary policy shocks on exchange rates. *Quarterly Journal of Economics*, 110:975–1009.

Engel, C. (2016). Exchange rates, interest rates, and the risk premium. *American Economic Review*, 106(2):436–474.

Ermolov, A. (2014). A unified theory of bond and currency markets. Working Paper.

Fama, E. (1984). Forward and spot exchange rates. *Journal of Monetary Economics*, 14(4):319–338.

Fama, E. and Bliss, R. (1987). The information in long-maturity forward rates. *American Economic Review*, 77:680–692.

Froot, K. (1989). New hope for the expectations hypothesis of the term structure of interest rates. *Journal of Finance*, 44:283–305.

Froot, K. and Frankel, J. (1989). Forward discount bias: Is it an exchange risk premium? *Quarterly Journal of Economics*, 104:139–161.

Giocoletti, M., Laursen, K., and Singleton, K. (2016). Learning, dispersion of beliefs and risk premiums in an arbitrage-free term structure model. Working Paper.

Gourinchas, P.-O. and Tornell, A. (2004). Exchange rate puzzles and distorted beliefs. *Journal of International Economics*, 64(2):303–333.

Greenwood, R. and Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies*, 27(3):714–746.

Hamilton, J. (1994). *Time Series Analysis*. Princeton University Press.

Harrison, J. and Kreps, D. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408.

Hassan, T. and Mano, R. (2017). Forward and spot exchange rates in a multi-currency world. Working Paper.

Hong, H. and Sraer, D. (2013). Quiet bubbles. *Journal of Financial Economics*, 110(3):596–606.

Ilut, C. (2012). Ambiguity aversion:implications for the uncovered interest parity puzzle. *American Economic Journal: Macroeconomics*, 4:33–65.

Jylhä, P. and Suominen, M. (2010). Speculative capital and currency carry trades. *Journal of Financial Economics*, 99(1).

Keynes, J. (1936). *The General Theory of Employment, Interest, and Money*. Palgrave Macmillan.

Koijen, R., Schmeling, M., and Vrugt, E. (2014). Survey expectations of returns and asset pricing puzzles. Working Paper.

Lustig, H., Roussanov, N., and Verdelhan, A. (2011). Common risk factors in currency markets. *Review of Financial Studies*, 24(11):3731–3777.

Lustig, H., Roussanov, N., and Verdelhan, A. (2014). Countercyclical currency risk premia. *Journal of Financial Economics*, 111(3):527–553.

Lustig, H., Stathopoulos, A., and Verdelhan, A. (2017). The term structure of currency carry trade risk premia. Working Paper.

McCallum, B. (1994). A reconsideration of the uncovered interest parity relationship. *Journal of Monetary Economics*, 33:105–132.

Moore, M. and Roche, M. (2002). Less of a puzzle: a new look at the forward forex market. *Journal of International Economics*, 58:387–411.

Newey, W. and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703?708.

Piazzesi, M., Salomao, J., and Schneider, M. (2015). Trend and cycle in bond premia. Working Paper.

Piazzesi, M. and Schneider, M. (2006). Equilibrium yield curves. *NBER Macro Annual*, (21):389–442.

Shiller, R. (1979). The volatility of long-term interest rates and expectations models of the term structure. *Journal of Political Economy*, 87:1190–1219.

Wachter, J. (2006). A consumption-based model of the term structure of interest rates. *Journal of Financial Economics*, 79:365–399.

Wright, J. (2011). Term premia and inflation uncertainty: Empirical evidence from an international panel dataset. *American Economic Review*, 101:1514–1534.

Xiong, W. and Yan, H. (2010). Heterogeneous expectations and bond markets. *Review of Financial Studies*, 23(4):1433–1466.

Zviadadze, I. (2016). Term structure of consumption risk premia in the cross section of currency returns. Journal of Finance, Forthcoming.

6 Appendix

6.1 Formulas Left Out in Proposition 1

The Kalman gains k_t, $\bar{k}_t$ and $\tilde{k}_t$ are given by

$$k_t = \frac{\lambda^2\sigma_t^2 + \sigma_\epsilon^2}{\lambda^2\sigma_t^2 + \sigma_\epsilon^2 + \sigma_v^2} \quad \bar{k}_t = \frac{\bar{\lambda}^2\bar{\sigma}_t^2 + \bar{\sigma}_\epsilon^2}{\bar{\lambda}^2\bar{\sigma}_t^2 + \bar{\sigma}_\epsilon^2 + \bar{\sigma}_v^2} \quad \tilde{k}_t = \frac{\lambda^2\sigma_t^2 + 2\sigma_\epsilon^2}{\lambda^2\sigma_t^2 + 2\sigma_\epsilon^2 + 2\sigma_v^2}. \tag{53}$$

The conditional volatilities of the persistent components are

$$\sigma_{t+1}^2 = (1-k_t)(\lambda^2\sigma_t^2 + \sigma_\epsilon^2) \qquad \bar{\sigma}_{t+1}^2 = (1-\bar{k}_t)(\bar{\lambda}^2\bar{\sigma}_t^2 + \bar{\sigma}_\epsilon^2) \tag{54}$$

for the first two states and the common state and

$$\tilde{\sigma}_{t+1}^2 = (1-\tilde{k}_t)(\lambda^2\tilde{\sigma}_t^2 + 2\sigma_\epsilon^2) \tag{55}$$

for the interest rate differential. The steady-state estimators are

$$\sigma^2 = \frac{1-k}{1-(1-k)\lambda^2}\sigma_\epsilon^2 \quad \bar{\sigma}^2 = \frac{1-\bar{k}}{1-(1-\bar{k})\bar{\lambda}^2}\bar{\sigma}_\epsilon^2 \quad \tilde{\sigma}^2 = 2\frac{1-k}{1-(1-k)\lambda^2}\sigma_\epsilon^2 \tag{56}$$

$$k = \tilde{k} = \frac{1 + \Delta - \eta(1 + \lambda^2)}{1 + \Delta + \eta(1 + \lambda^2)} \quad \bar{k} = \frac{1 + \bar{\Delta} - \bar{\eta}(1 + \bar{\lambda}^2)}{1 + \bar{\Delta} + \bar{\eta}(1 + \bar{\lambda}^2)}. \tag{57}$$

Here

$$\Delta^2 = [\eta(1 - \lambda^2) + 1]^2 + 4\eta\lambda^2 \quad \bar{\Delta}^2 = [\bar{\eta}(1 - \bar{\lambda}^2) + 1]^2 + 4\bar{\eta}\bar{\lambda}^2 \tag{58}$$

and

$$\eta = \frac{\sigma_v^2}{\sigma_\epsilon^2} \quad \bar{\eta} = \frac{\bar{\sigma}_v^2}{\bar{\sigma}_\epsilon^2}. \tag{59}$$

6.2 Proof of Proposition 2

The expression for $A(1)$ and $\mathbf{B}(1)$ follow immediately. Given the conjecture, log bond prices and SDF are (conditionally) jointly normal. To see the recursion formulas note

$$q_t(n) = \mathbb{E}_t^S[m_{t+1} + q_{t+1}(n-1)] + \frac{1}{2}\mathbb{V}ar_t^S[m_{t+1} + q_{t+1}(n-1)] \tag{60}$$

or

$$A(n) + \mathbf{B}(n)'\mathbf{Y}_t = \mathbb{E}_t^S[m_{t+1} + A(n-1) + \mathbf{B}(n-1)'\mathbf{Y}_{t+1}] + \tag{61}$$

$$\frac{1}{2}\mathbb{V}ar_t^S[m_{t+1} + A(n-1) + \mathbf{B}(n-1)'\mathbf{Y}_{t+1}]. \tag{62}$$

Note that

$$\mathbb{E}_t^S[m_{t+1} + A(n-1) + \mathbf{B}(n-1)'\mathbf{Y}_{t+1}] = -logR - \frac{\bar{\sigma}_\epsilon^2\bar{\varphi}_t^2}{2} - \frac{\sigma_\epsilon^2\varphi_t^2}{2} - z_t - \bar{z}_t + \tag{63}$$

$$A(n-1) + \mathbf{B}(n-1)'[\mathbb{E}_t^S[z_{t+1}], \mathbb{E}_t^S[\bar{z}_{t+1}], \mathbb{E}_t^S[z_{t+2}], \mathbb{E}_t^S[\bar{z}_{t+2}]]. \tag{64}$$

Also

$$\mathbb{E}_t^S[z_{t+2}] = \lambda\mathbb{E}_t^S[z_{t+1}] \qquad \mathbb{E}_t^S[\bar{z}_{t+2}] = \bar{\lambda}\mathbb{E}_t^S[\bar{z}_{t+1}]. \tag{65}$$

Therefore after some algebra

$$\mathbb{E}^S_{t+1}[m_{t+1} + A(n-1) + \mathbf{B}(n-1)'\mathbf{Y}_{t+1}] = -logR - \frac{\bar{\sigma}^2_\epsilon \bar{\varphi}^2_t}{2} - \frac{\sigma^2_\epsilon \varphi^2_t}{2} - z_t - \bar{z}_t \quad (66)$$

$$+A(n-1) + \mathbb{E}^S_t[z_{t+1}](B_1(n-1) + \lambda B_3(n-1)) + \quad (67)$$

$$\mathbb{E}^S_t[\bar{z}_{t+1}](B_2(n-1) + \bar{\lambda} B_4(n-1)). \quad (68)$$

The variance is given by (leaving out some algebra)

$$\mathbb{V}ar^S_t[m_{t+1} + A(n-1) + \mathbf{B}(n-1)'\mathbf{Y}_{t+1}] = \quad (69)$$

$$[\varphi^2_t - 2\varphi_t(B_1(n-1) + k\lambda B_3(n-1))]\sigma^2_\epsilon + [\bar{\varphi}^2_t - 2\bar{\varphi}_t(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))]\bar{\sigma}^2_\epsilon \quad (70)$$

$$+(\sigma^2 + \sigma^2_v)(B_1(n-1) + k\lambda B_3(n-1))^2 + (\bar{\sigma}^2 + \bar{\sigma}^2_v)(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))^2. \quad (71)$$

$$(72)$$

Therefore I obtain the equation

$$A(n) + B_1(n)z_t + B_2(n)\bar{z}_t + B_3(n)\mathbb{E}^s_t[z_{t+1}] + B_4(n)\mathbb{E}^s_t[\bar{z}_{t+1}] = \quad (73)$$

$$-logR - \frac{\bar{\sigma}^2_\epsilon \bar{\varphi}^2_t}{2} - \frac{\sigma^2_\epsilon \varphi^2_t}{2} - z_t - \bar{z}_t + A(n-1) + \mathbb{E}^s_t[z_{t+1}](B_1(n-1) + \lambda B_3(n-1)) + \quad (74)$$

$$\mathbb{E}^s_t[\bar{z}_{t+1}](B_2(n-1) + \bar{\lambda}B_4(n-1)) + \quad (75)$$

$$\frac{1}{2}[\varphi^2_t - 2\varphi_t(B_1(n-1) + k\lambda B_3(n-1))]\sigma^2_\epsilon + \frac{1}{2}[\bar{\varphi}^2_t - 2\bar{\varphi}_t(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))]\bar{\sigma}^2_\epsilon \quad (76)$$

$$+\frac{1}{2}(\sigma^2 + \sigma^2_v)(B_1(n-1) + k\lambda B_3(n-1))^2 + \frac{1}{2}(\bar{\sigma}^2 + \bar{\sigma}^2_v)(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))^2. \quad (77)$$

From this one can solve

$$A(n) = -logR + A(n-1) + \quad (78)$$

$$-\varphi_0(B_1(n-1) + k\lambda B_3(n-1))\sigma^2_\epsilon - \bar{\varphi}_0(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))\bar{\sigma}^2_\epsilon \quad (79)$$

$$+\frac{1}{2}(\sigma^2 + \sigma^2_v)(B_1(n-1) + k\lambda B_3(n-1))^2 + \frac{1}{2}(\bar{\sigma}^2 + \bar{\sigma}^2_v)(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))^2 \quad (80)$$

$$B_1(n) = -1 - \varphi_1(B_1(n-1) + k\lambda B_3(n-1))\sigma_\epsilon^2 - \bar{\varphi}_1(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))\bar{\sigma}_\epsilon^2 \tag{81}$$

$$B_2(n) = -1 \tag{82}$$

$$B_3(n) = \lambda B_3(n-1) + B_1(n-1) \quad B_4(n) = \bar{\lambda}B_4(n-1) + B_2(n-1). \tag{83}$$

The coefficient values in the simplified case with non-varying market prices of risk follow easily. Here one has

$$A(n) = -logR + A(n-1)+ \tag{84}$$

$$-\varphi_0(B_1(n-1) + k\lambda B_3(n-1))\sigma_\epsilon^2 - \bar{\varphi}_0(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))\bar{\sigma}_\epsilon^2 \tag{85}$$

$$+\frac{1}{2}(\sigma^2 + \sigma_v^2)(B_1(n-1) + k\lambda B_3(n-1))^2 + \frac{1}{2}(\bar{\sigma}^2 + \bar{\sigma}_v^2)(B_2(n-1) + \bar{k}\bar{\lambda}B_4(n-1))^2. \tag{86}$$

6.3 Proof of Proposition 3

The standard complete market condition in logs is

$$m_{t+1} + s_{t+1} - s_t = m_{t+1}^* \tag{87}$$

Taking expectations

$$s_t = \mathbb{E}_t^S[m_{t+1} - m_{t+1}^*] + \mathbb{E}_t^S[s_{t+1}] \tag{88}$$

or

$$s_t = x_t - \frac{\bar{\sigma}_\epsilon^2\bar{\varphi}_t^2}{2} - \frac{\sigma_\epsilon^2\varphi_t^2}{2} + \frac{\bar{\sigma}_\epsilon^2\bar{\varphi}_t^{*2}}{2} + \frac{\sigma_\epsilon^2\varphi_t^{*2}}{2} + \mathbb{E}_t^S[s_{t+1}]. \tag{89}$$

Iterating forward one obtains

$$s_t = \mathbb{E}_t^S\sum_{j=0}^{\infty}x_{t+j} + \mathbb{E}_t^S\sum_{j=0}^{\infty}\Gamma_{t+j} + \bar{s}_t, \tag{90}$$

where

$$\Gamma_t = -\frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^2}{2} - \frac{\sigma_\epsilon^2 \varphi_t^2}{2} + \frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^{*2}}{2} + \frac{\sigma_\epsilon^2 \varphi_t^{*2}}{2} \tag{91}$$

and $\bar{s}_t = \underset{j\to\infty}{lim} \mathbb{E}_t^S[s_{t+j}]$ is assumed to be well-defined. As in Gourinchas and Tornell (2004) I assume the agents hold correct views concerning this component. Therefore it does not affect return predictability and will be ignored in the calculations. The sums converge.

Furthermore

$$\mathbb{E}_t^S \sum_{j=0}^{\infty} x_{t+j} = x_t + \frac{1}{1-\lambda} \mathbb{E}_t^S[x_{t+1}]. \tag{92}$$

Therefore (leaving out some algebra)

$$\mathbb{E}_t \left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} x_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{\infty} x_{t+j} \right] = -x_t + \left[1 + \frac{\lambda k}{1-\lambda} \right] \left[\mathbb{E}_t x_{t+1} - \mathbb{E}_t^S x_{t+1} \right]. \tag{93}$$

(94)

Furthermore

$$\Gamma_t = -\frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^2}{2} - \frac{\sigma_\epsilon^2 \varphi_t^2}{2} + \frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^{*2}}{2} + \frac{\sigma_\epsilon^2 \varphi_t^{*2}}{2} = \tag{95}$$

$$-\frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_1^2}{2} z_t^2 - \bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 z_t - \frac{\sigma_\epsilon^2 \varphi_1^2}{2} z_t^2 - \sigma_\epsilon^2 \varphi_0 \varphi_1 z_t + \tag{96}$$

$$\frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_1^2}{2} z_t^{*2} + \bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 z_t^* + \frac{\sigma_\epsilon^2 \varphi_1^2}{2} z_t^{*2} + \sigma_\epsilon^2 \varphi_0 \varphi_1 z_t^*. \tag{97}$$

(98)

Typically the squared terms are small relative to the other terms. One can write

$$\Theta_t = \left[1 + \frac{\lambda k}{1-\lambda} \right] \left[\mathbb{E}_t x_{t+1} - \mathbb{E}_t^S x_{t+1} \right] + \mathbb{E}_t \left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} \Gamma_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{\infty} \Gamma_{t+j} \right]. \tag{99}$$

(100)

Note that assuming $\varphi_1 = \bar{\varphi}_1 = 0$ means the second term equals zero. Next one needs to derive an expression for the relative long-term bond price. Note

$$q_t^*(n) - q_t(n) = B_1(n)(z_t^* - z_t) + B_3(n)(\mathbb{E}_t^S[z_{t+1}^*] - \mathbb{E}_t^S[z_{t+1}]) = \tag{101}$$
$$B_1(n)x_t + B_3(n)(\mathbb{E}_t^S[z_{t+1}^*] - \mathbb{E}_t^S[z_{t+1}]) \tag{102}$$

and hence

$$\mathbb{E}_t[q_t^*(n-1) - q_t(n-1)] - [q_t^*(n) - q_t(n)] = \tag{103}$$
$$B_1(n-1)\lambda x_t - B_1(n)x_t + B_3(n-1)\mathbb{E}_t[\mathbb{E}_{t+1}^S[z_{t+2}] - \mathbb{E}_t^S[z_{t+2}^*]] \tag{104}$$
$$-B_3(n)[\mathbb{E}_t^S[z_{t+1}] - \mathbb{E}_t^S[z_{t+1}^*]]. \tag{105}$$

Therefore

$$\Theta_t^n = \mathbb{E}_t[q_t^*(n-1) - q_t(n-1) + s_{t+1}] - [q_t^*(n) - q_t(n)] - s_t = \tag{106}$$
$$B_1(n-1)\lambda x_t - B_1(n)x_t + B_3(n-1)\mathbb{E}_t[\mathbb{E}_{t+1}^S[z_{t+2}] - \mathbb{E}_t^S[z_{t+2}^*]] \tag{107}$$
$$-B_3(n)[\mathbb{E}_t^S[z_{t+1}] - \mathbb{E}_t^S[z_{t+1}^*]] - x_t + \left[1 + \frac{\lambda k}{1-\lambda}\right]\left[\mathbb{E}_t x_{t+1} - \mathbb{E}_t^S x_{t+1}\right]. \tag{108}$$

Finally I derive a simple expression for the above expected return assuming $\varphi_1 = \bar{\varphi}_1 = 0$. Due to log-normality

$$q_t(n) = log(P_t^n) = \mathbb{E}_t^S[\sum_{i=1}^{n} m_{t+i-1,t+i}] + \frac{1}{2}\mathbb{V}ar_t^S[\sum_{i=1}^{n} m_{t+i-1,t+i}]. \tag{109}$$

Variance terms for the two countries are now equal and therefore

$$\mathbb{E}_t^S[\sum_{i=1}^{n}(m_{t+i-1,t+i}^* - m_{t+i-1,t+i})] = -\mathbb{E}_t^S \sum_{s=t}^{n} x_s. \tag{110}$$

Now one can solve

$$\mathbb{E}_t^S[\sum_{i=1}^{n} m_{t+i-1,t+i}] = -x_t - \frac{1-\lambda^n}{1-\lambda}\mathbb{E}_t^S[x_{t+1}]. \tag{111}$$

Therefore (leaving out some algebra)

$$\mathbb{E}_t[q^*_{t+1}(n-1) - q_{t+1}(n-1)] - [q^*_t(n) - q_t(n)] = \tag{112}$$

$$x_t - \mathbb{E}_t[x_{t+1}]\left[1 + k\lambda\frac{1-\lambda^{n-1}}{1-\lambda}\right] + \mathbb{E}^S_t[x_{t+1}]\left[\frac{1-\lambda^n}{1-\lambda} - (1-k)\lambda\frac{1-\lambda^{n-1}}{1-\lambda}\right]. \tag{113}$$

$$\tag{114}$$

I obtain (leaving out some algebra)

$$\Theta^n_t = \left[\mathbb{E}_t[x_{t+1}] - \mathbb{E}^S_t[x_{t+1}]\right]\frac{k\lambda^n}{1-\lambda}. \tag{115}$$

6.4 Proof of Proposition 4

I look at the covariance with the conditional expectation function

$$\delta_1 = \frac{cov(x_t, \Theta_t)}{Var(x_t)}. \tag{116}$$

Here

$$cov(x_t, \Theta_t) = \left[1 + \frac{\lambda k}{1-\lambda}\right] cov(x_t, \mathbb{E}_t[x_{t+1}] - \mathbb{E}^S_t[x_{t+1}]) = \tag{117}$$

$$Var(x_t)\lambda\left[1 + \frac{\lambda k}{1-\lambda}\right] - \left[1 + \frac{\lambda k}{1-\lambda}\right] cov(x_t, \mathbb{E}^S_t[x_{t+1}]). \tag{118}$$

Also

$$\mathbb{E}^S_t[x_{t+1}] = k\lambda x_t + k(1-k)\lambda^2 x_{t-1} + k(1-k)^2\lambda^3 x_{t-2} + \ldots \tag{119}$$

$$cov(x_t, \mathbb{E}^S_t[x_{t+1}]) = Var(x_t)[\lambda k + k(1-k)\lambda^3 + k(1-k)^2\lambda^5 + \ldots] \tag{120}$$

$$= \frac{\lambda k}{1-(1-k)\lambda^2} Var(x_t). \tag{121}$$

Hence

$$cov(x_t, \Theta_t) = -Var(x_t)\left[1 + \frac{\lambda k}{1-\lambda}\right]\left[\frac{\lambda k - \lambda + (1-k)\lambda^3}{1-(1-k)\lambda^2}\right] \quad (122)$$

and

$$\delta_1 = -\left[1 + \frac{\lambda k}{1-\lambda}\right]\left[\frac{\lambda k - \lambda + (1-k)\lambda^3}{1-(1-k)\lambda^2}\right]. \quad (123)$$

Also

$$\delta_1^n = -\frac{k\lambda^n}{1-\lambda}\left[\frac{\lambda k - \lambda + (1-k)\lambda^3}{1-(1-k)\lambda^2}\right]. \quad (124)$$

One can see that the expression is positive. The other results follow easily.
□

6.5 Proof of Proposition 5

Consider the regression

$$s_{t+j} - \mathbb{E}_t^S[s_{t+j}] = \kappa_{10} + \kappa_{11} x_t + e_{t+j}. \quad (125)$$

One can write

$$s_{t+j} - \mathbb{E}_t^S[s_{t+j}] = \sum_{k=0}^{\infty} \mathbb{E}_{t+j}^S[x_{t+j+k}] - \sum_{k=0}^{\infty} \mathbb{E}_t^S[x_{t+j+k}] + \bar{s}_{t+j} - \mathbb{E}_t^S[\bar{s}_{t+j}]. \quad (126)$$

Because I assumed $\mathbb{E}_t^S[\bar{s}_{t+j}] = \mathbb{E}_t[\bar{s}_{t+j}]$ the permanent component does not affect the results. One needs to evaluate

$$Cov(\mathbb{E}_{t+j}^S[x_{t+j+k}] - \mathbb{E}_t^S[x_{t+j+k}], x_t) = \quad (127)$$

$$\lambda^{k-1} Cov(\mathbb{E}_{t+j}^S[x_{t+j+1}] - \lambda^j \mathbb{E}_t^S[x_{t+1}], x_t). \quad (128)$$

Recall that

$$\mathbb{E}^S_{t+j}[x_{t+j+1}] = (1-k)\lambda\mathbb{E}^S_{t+j-1}[x_{t+j}] + k\lambda x_{t+j} = \tag{129}$$

$$(1-k)^j\lambda^j\mathbb{E}^S_t[x_{t+1}] + k\lambda[x_{t+j} + \lambda(1-k)x_{t+j-1} + \ldots + \lambda^j(1-k)^j x_t]. \tag{130}$$

Hence after some algebra

$$Cov(\mathbb{E}^S_{t+j}[x_{t+j+1}] - \lambda^j\mathbb{E}^S_t[x_{t+1}], x_t) = \tag{131}$$

$$[(1-k)^j - 1]\lambda^j Cov(\mathbb{E}^S_t[x_{t+1}], x_t) + \lambda^{j+1}[1-(1-k)^{j+1}]Var(x_t). \tag{132}$$

$$\tag{133}$$

On the other hand (see the proof of proposition 4)

$$Cov(\mathbb{E}^S_t[x_{t+1}], x_t) = \frac{\lambda k}{1-(1-k)\lambda^2}Var(x_t). \tag{134}$$

Therefore

$$Cov(\mathbb{E}^S_{t+j}[x_{t+j+1}] - \lambda^j\mathbb{E}^S_t[x_{t+1}], x_t) = \lambda^{j+1}\left[\frac{k(1-k)^j - k}{1-(1-k)\lambda^2} + 1 - (1-k)^{j+1}\right]Var(x_t). \tag{135}$$

The term

$$\left[\frac{k(1-k)^j - k}{1-(1-k)\lambda^2} + 1 - (1-k)^{j+1}\right] \tag{136}$$

governs the sign of κ_{11} in the the longer horizon FX rate regression. Now

$$\left[\frac{k(1-k)^j - k}{1-(1-k)\lambda^2} + 1 - (1-k)^{j+1}\right] > (1-k)^j - (1-k)^{j+1} > 0. \tag{137}$$

Hence $\kappa_{11} > 0$.

□

6.6 Yield Volatilities

Variance of the yield of maturity n is given by

$$\mathbb{V}ar\left(\frac{12}{n}q(n)\right) = \frac{12^2}{n^2}\mathbf{B}(n)'\Sigma\mathbf{B}(n), \tag{138}$$

where

$$\Sigma = \mathbb{V}ar(\mathbf{Y}_t) \tag{139}$$

and $\mathbf{Y}_t = [\hat{z}_t, \mathbb{E}_t^s[z_{t+1}], \mathbb{E}_t^s[\bar{z}_{t+1}]]'$. Furthermore

$$\mathbb{V}ar(\hat{z}_t) = Var(z_t) + Var(\bar{z}_t) \tag{140}$$

$$\mathbb{V}ar(\mathbb{E}_t^S[z_{t+1}]) = \left[\frac{k^2\lambda^2}{1-(1-k)^2\lambda^2} + \frac{2(1-k)}{1-(1-k)^2\lambda^2}\frac{\lambda^2 k}{1-(1-k)\lambda^2}\right]Var(z_t) \tag{141}$$

$$\mathbb{V}ar(\mathbb{E}_t^S[\bar{z}_{t+1}]) = \left[\frac{\bar{k}^2\bar{\lambda}^2}{1-(1-\bar{k})^2\bar{\lambda}^2} + \frac{2(1-\bar{k})}{1-(1-\bar{k})^2\bar{\lambda}^2}\frac{\bar{\lambda}^2 \bar{k}}{1-(1-\bar{k})\bar{\lambda}^2}\right]Var(\bar{z}_t) \tag{142}$$

$$\mathbb{C}ov(\mathbb{E}_t^S[\bar{z}_{t+1}], z_t) = 0 \qquad \mathbb{C}ov(\mathbb{E}_t^S[z_{t+1}], \bar{z}_t) = 0 \tag{143}$$

$$\mathbb{C}ov(\mathbb{E}_t^S[z_{t+1}], z_t) = \frac{\lambda k}{1-(1-k)\lambda^2}\mathbb{V}ar(z_t) \quad \mathbb{C}ov(\mathbb{E}_t^S[\bar{z}_{t+1}], \bar{z}_t) = \frac{\bar{\lambda}\bar{k}}{1-(1-\bar{k})\bar{\lambda}^2}\mathbb{V}ar(\bar{z}_t). \tag{144}$$

6.7 The Stochastic Discount Factor Under the Objective Measure

This section shows how to rewrite the home stochastic discount factor under the objective measure. It is seen that the stochastic discount factor is strictly positive under the objective measure. This implies that the economy does not allow for arbitrage opportunities under either the subjective nor objective measure. Put alternatively, the objective and subjective probability measures are equivalent.

The time t price of a payoff X_{t+1} is given by

$$P_t = \mathbb{E}_t^S[M_{t,t+1}X_{t+1}]. \tag{145}$$

This can be rewritten under the objective measure as

$$P_t = \mathbb{E}_t[M_{t,t+1}\xi_{t,t+1}X_{t+1}], \tag{146}$$

where $\xi_{t,t+1} = \frac{dS}{dP}$ is the Radon-Nikodym derivative and $M_{t,t+1}\xi_{t,t+1}$ is the SDF under the objective measure. Here one can see that the objective stochastic discount factor is obtained by perturbing the subjective discount factor by the Radon-Nikodym derivative. This puts more weight on events that are relatively more probable under the subjective than the objective measure.

A natural state variable in the economy is $(z_t, z_t^*, \bar{z}_t, \mathbb{E}_t^S[z_{t+1}], \mathbb{E}_t^S[z_{t+1}^*], \mathbb{E}_t^S[\bar{z}_{t+1}], t) \equiv (\mathbf{Z}_t, t)$, where t affects the economy through the updating coefficients and variances of persistent components. Let us therefore assume that X_{t+1} depends only on these state variables as well as the shocks, that is, $X_{t+1} \equiv X_{t+1}(\mathbf{Z}_{t+1}, \boldsymbol{\epsilon}_{t+1})$. Let us define $\mathcal{H}_{t+1} \equiv (\boldsymbol{\epsilon_{t+1}}, \mathbf{Z}_{t+1})$. $\mathcal{H}_{t+1}$ is Gaussian both under the subjective and objective measure.

Now one can write

$$P_t(\mathbf{Z}_t) = \mathbb{E}[M_{t,t+1}\xi_{t,t+1}(\mathcal{H}_{t+1}|\mathbf{Z}_t)X_{t+1}(\mathcal{H}_{t+1})|\mathbf{Z}_t, t]. \tag{147}$$

Because of the Gaussian nature of the state variables and shocks

$$\xi_{t,t+1}(\mathcal{H}_{t+1}|\mathbf{Z}_t) = \frac{\sqrt{|\Sigma_t|}}{\sqrt{|\Sigma_t^S|}} exp\left(-\frac{1}{2}\phi_{t,t+1}^S(\mathcal{H}_{t+1}|\mathbf{Z}_t) + \frac{1}{2}\phi_{t,t+1}(\mathcal{H}_{t+1}|\mathbf{Z}_t)\right), \tag{148}$$

where the terms inside the brackets are

$$\phi_{t,t+1}^S(\mathcal{H}_{t+1}|\mathbf{Z}_t) = (\mathcal{H}_{t+1} - \mathbb{E}^S[\mathcal{H}_{t+1}|\mathbf{Z}_t])'(\Sigma_t^S)^{-1}(\mathcal{H}_{t+1} - \mathbb{E}^S[\mathcal{H}_{t+1}|\mathbf{Z}_t]). \tag{149}$$

and

$$\phi_{t,t+1}(\mathcal{H}_{t+1}|\mathbf{Z}_t) = (\mathcal{H}_{t+1} - \mathbb{E}[\mathcal{H}_{t+1}|\mathbf{Z}_t])'(\Sigma_t)^{-1}(\mathcal{H}_{t+1} - \mathbb{E}[\mathcal{H}_{t+1}|\mathbf{Z}_t]). \tag{150}$$

Moreover, the covariances matrices are given by

$$\Sigma_t^S = Var_t^S(\mathcal{H}_{t+1}|\mathbf{Z}_t) \tag{151}$$

$$\Sigma_t = Var_t(\mathcal{H}_{t+1}|\mathbf{Z}_t). \tag{152}$$

These could be calculated similarly to the variance and covariance terms for the yield volatility calculations. Furthermore, the expectation terms can be solved using the updating formulas for the Kalman filter and the correct processes for the state variables (the shocks have zero expectation under both measures). Note that because $\Sigma_t^S \neq \Sigma_t$, the Radon-Nikodym derivative between the subjective and objective measure is not of the form assumed for example by Piazzesi et al. (2015).

By computing Σ_t^S and Σ_t one can see that these covariance matrices are not degenerate in that no state variable or shock would be perfectly correlated with another state variable or shock or that some shock or state variable had zero variance. This implies that the determinants are non-zero as well as that the inverse covariance matrices are well-defined. Therefore the Radon-Nikodym derivative $\xi_{t,t+1}$ is well defined and strictly positive. Therefore the SDF is strictly positive under the objective measure.

A well-known result in asset pricing is that the existence of a strictly positive stochastic discount factor (or equivalent martingale measure) is equivalent to no-arbitrage (Harrison and Kreps, 1979). Therefore the economy does not allow for arbitrage opportunities under the objective (or subjective) measure. An analogous derivation could be performed for the foreign stochastic discount factor.