Bonds, Currencies and Expectational Errors

Eleonora Granziera and Markus Sihvonen*

November 24, 2020

Abstract

We propose a model in which sticky expectations concerning short-term interest rates generate joint predictability patterns in bond and currency markets. Using our calibrated model, we quantify the effect of this channel and find that it largely explains why short rates and yield spreads predict bond and currency returns. The model also creates the downward sloping term structure of carry trade returns documented by Lustig et al. (2019), difficult to replicate in a rational expectations framework. Consistent with the model, we find that variables that predict bond and currency returns also predict survey-based expectational errors concerning interest and FX rates. Including a sticky short rate expectations channel into a standard affine term structure model improves its fit and allows the model to better capture the drift patterns in the data.

Keywords: Bond and currency premia, sticky expectations, interest rate forecast errors

JEL classification: E43, F31, D84

*Norges Bank, Bankplassen 2, P.O. Box 1179, Sentrum 0107, Oslo, Norway; Bank of Finland, Research Unit, Snellmaninaukio, P.O. Box 160, 00101 Helsinki, Finland. Email: eleonora.granziera@gmail.com; mesihvonen@gmail.com (corresponding author). The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Norges Bank, the Bank of Finland or the Eurosystem. We thank Anna Cieślak, Olivier Coibion, Pierre Collin-Dufresne, Andrey Ermolov, Pierre-Olivier Gourinchas, Simas Kučinskas, Kevin Lansing, George Pennacchi, Matti Suominen, an anonymous referee from the Norges Bank Working Paper series and seminar participants at the Bank of Finland Research Seminar, HEC Paris Finance PhD Workshop and Econometric Society European Winter Meetings for useful comments. We thank Janne Lehto and Ilona Vänni for excellent research assistance.

1 Introduction

This paper presents the first unified theory of bond and currency markets based on expectational errors. According to this theory forecast errors concerning short-term interest rates give rise to joint predictability patterns in bond and currency markets. These predictability patterns nest, and can explain, many of the predictability puzzles documented in the previous literature.

Lustig et al. (2019) argue that the literature's key findings concerning currency and bond return predictability are related: while a high short-term interest rate predicts high returns for a currency, it predicts low returns for long-term bonds denominated in this currency. Similarly, a steep slope of the yield curve predicts low returns for a currency but high returns for corresponding long-term bonds. Such negative correlation between the currency and bond premia represents a puzzle for rational expectations macrofinance models. The model presented in this paper explains this correlation.

Our model is based on the well-documented finding that forecasters update their short rate predictions sluggishly (Coibion and Gorodnichenko, 2015). We do not offer an explanation for this pattern, though we note that it can be caused indirectly due to slow updating concerning factors driving interest rates.[1] However, the key assumption of our approach is that currencies and bonds are priced consistently with such biased expectations concerning short rates.

Then, the return on a bond or currency can be decomposed into a rational risk premium, a short rate misperception effect and a risk premium misperception effect. This decomposition is an identity, it holds in all models in which subjective expectations are given by a probability measure. We argue that under relatively weak and realistic conditions the contribution of short rate forecast errors to return variation can be identified econometrically.

We use our calibrated model to quantify the effect of the interest rate misperception channel. We find that it can account for most of the variation in bond and currency premia driven by changes in short rates and yield spreads. The channel generates coefficients in predictability regressions similar to those found in the data.

Various authors, including Gourinchas and Tornell (2004), Cieslak (2017)

[1] There are various possible explanations, for example D'Acunto et al. (2019) argue that household forecast errors are related to *cognitive frictions*. Sticky expectations are also consistent with *inattention* (see e.g. Gabaix (2019)). Moreover, Ilut (2012) notes that similar effects follow from models with ambiguity averse preferences.

and Piazzesi et al. (2015), have explored the effects of expectational errors on bond and currency returns separately. However, what has heretofore been unnoticed is that expectational errors concerning short rates provide a natural candidate for a *joint* theory of bond and currency markets.

The economic intuition behind our key results is simple. The current home and foreign short-term interest rates are known but agents must forecast their future values. The value of a foreign currency is increasing in expected foreign short-term interest rates and the value of foreign long-term bond decreasing in expected (foreign) short-term interest rates. When agents underpredict the path of future foreign interest rates, the value of the foreign currency is lower than under rational expectations but the value of the foreign bond higher than under rational expectations. This implies high actual returns for the currency but low returns for the corresponding bond.

In the data this underprediction is associated with *sticky expectations*. When short-term interest rates increase, for example due to a contractionary monetary policy shock, it takes time for forecasters to revise their future short rate expectations up. This leads forecasters to underpredict the future path of short rates. As the forecasters slowly increase their expectations over future foreign short-term interest rates, the foreign currency appreciates but the value of the foreign bond falls. Before the forecasters have updated their expectations closer to rational values, the returns for a currency will be high but the returns for the bond low.

Note that sticky expectations gives rise to a relation between the *level* of short-term interest rates and the degree of underprediction concerning future interest rates. When short-term interest rates are high, they have on average increased recently. Therefore high short-term interest rates are associated with larger underprediction concerning future interest rates. This implies that a high short-term interest rate predicts high returns for a currency but low returns for the corresponding long-term bond.

We now demonstrate this intuition further with a simplified version of the model. Assume that the currencies are subject to similar perceived risk premia. Denote the short-term interest rate differential between the foreign and home country by $x_t \equiv i_t^* - i_t$ and the log FX rate by s_t, where an increase in s_t implies an appreciation of the foreign currency. 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).[2] 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 under subjective expectations the interest rate differential tends to remain lower than under rational expectations $\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}] > 0$. 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 $\mathbb{E}_t[s_{t+1}] - \mathbb{E}_t^S[s_{t+1}] > 0$. That is, the foreign currency will be stronger on average the next period than predicted by forecasters.

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

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

When $x_t > 0$ the price of the foreign bond, $q_t^*(n)$, that is known by all agents, is relatively low and the yield high. However, because this is due to a recent interest rate shock the forecasters believe $\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}] > 0$ 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, measured under rational expectations, as high interest rate currencies have long-term bonds that are overpriced compared to prices under rational expectations.

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. At the same time the long-term bond of the same currency is relatively overpriced and yields low actual returns. Higher maturity increases the sensitivity of a bond to predictions about future interest rates, so this effect is stronger the longer the maturity of the bond. 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 domestic currency returns. This explains why the term structure of expected carry trade returns is downward sloping.

[2] We discuss the role of the permanent component of the FX rate later.

We provide strong empirical evidence that supports the importance of short rate forecast errors for bond and currency returns. In particular, we show that the same variables that predict bond and currency returns also predict survey-based expectational errors concerning FX rates and long-term interest rates. For example, when (domestic or foreign) short-term interest rates are high, forecasters underestimate the future level of long-term interest rates and overestimate the future value of long-term bonds. Similarly, when foreign short-term interest rates are high relative to domestic interest rates, forecasters underestimate the future value of the foreign currency relative to the home currency. Moreover, we show that foreign currency returns tend to be particularly high, and bond returns low, when foreign short rates have recently increased.

Finally, we discuss the results in the context of affine term structure models. We show that such models can be amended to incorporate sticky short rate expectations. The sticky expectations version of a standard affine term structure model is both consistent with survey data but also gives a more acccurate match to the predictability patterns in the data.

Related Literature This paper contributes to the vast literature on markets for currencies and government bonds. Special attention is given to explaining predictability patterns in bond and currency returns. The seminal paper for currencies is Fama (1984) which finds that currencies with high short-term interest rates appreciate rather than depreciate as predicted by uncovered interest rate parity. On the other hand, Fama and Bliss (1987) and Cochrane and Piazzesi (2005) find that high bond yields are associated with high bond returns, a violation of the expectations hypothesis. Lustig et al. (2019) argue that these two findings are related as high relative bond yields predict low returns for the corresponding currency.

A large literature in the tradition of rational expectations consumption based asset pricing has attempted to explain the predictability patterns in bond and currency markets. Examples include applications of the habit model for the bond market (see e.g. Wachter (2006)) and those for the currency market (see e.g. Verdelhan (2010)). Moreover, e.g. Bansal and Shaliastovich (2012) apply the long-run risk model for both bonds and currencies.

A second literature in the tradition of no-arbitrage term structure models (see e.g. Duffie and Kan (1996)) has taken a more reduced form approach to modeling bonds. Similar models have been applied to currencies (see e.g. Backus et al. (2001) and Lustig et al. (2011)). Moreover, e.g. Sarno et al. (2012) studies the joint performance of a four factor affine model in pricing bonds and currencies. Note that Lustig et al. (2019) argue that neither the standard structural models nor these no-arbitrage models are able to replicate the term structure of carry trade returns.

A key alternative to the risk-based approach is to relax the assumption of rational expectations. This choice can be motivated by the systematic expectational errors documented in surveys (see e.g. Bacchetta et al. (2009), Coibion and Gorodnichenko (2012) and Greenwood and Shleifer (2014)). 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 (2017). However, to our best knowledge this is the first paper that offers a joint explanation for bond and currency markets based on expectational errors.

The above mentioned risk-based models are based on the assumption of frictionless markets. Jylhä and Suominen (2010) and Gabaix and Maggiori (2015) argue that financial frictions can explain currency carry trade returns. In concurrent work Greenwood et al. (2019) posit that asset market frictions can explain both the properties of bonds and currencies, including the downward sloping term structure of carry trade returns. These effects can potentially complement those presented in this paper.

2 Determinants of Bond and Currency Premia

2.1 A General Framework

We first introduce the general model structure. Let there be two probability measures P and S. Here P corresponds to objective probabilities as viewed by a rational econometrician. On the other hand, S represents the subjective beliefs of an investor. The standard assumption in the literature is that $P = S$, but this paper argues that is better to view these two as separate. If different investors have heterogeneous beliefs, we can also define S as a weighted average of the individual probabilities of the agents.[3] For simplicity we omit the P-symbol from expectations taken under rational beliefs. To rule out ill-defined cases we assume the probability measures are equivalent, that is

[3] We do not explicitly address disagreement. For the effects of disagreement on bond markets see e.g. Buraschi and Whelan (2012) and Giacoletti et al. (2018), for FX markets see Buraschi et al. (2010).

they agree on zero probability events.

Without loss of generality, consider the case of two countries, home and foreign, where the latter variables are denoted by stars. As in the introduction, the home and foreign log prices of nominal zero coupon bonds of maturities n are $q_t(n)$ and $q_t(n)^*$. Moreover, the short rate difference between the two countries is $x_t \equiv i_t^* - i_t$ and the log nominal exchange rate is s_t.

We are particularly interested in explaining the returns from currency carry trades implemented with different maturity bonds. Similarly to Lustig et al. (2019) we define the logarithmic returns from the standard carry trade and maturity n bond carry trade as

$$r_{t+1}^{FX} \equiv x_t + s_{t+1} - s_t \tag{4}$$

$$r_{t+1}^{FX}(n) \equiv q_{t+1}^*(n-1) - q_{t+1}(n-1) - (q_t^*(n) - q_t(n)) + s_{t+1} - s_t \tag{5}$$

Here the former corresponds to a return from a strategy of investing in foreign currency short term bills and borrowing short term in the home currency. The latter gives the return from the same trade but implemented by buying a foreign currency n maturity bond and selling short a home currency n maturity bond. We can also define the relative excess return from n maturity zero-coupon bonds as

$$r_{t+1}^{B}(n) \equiv r_{t+1}^{FX}(n) - r_{t+1}^{FX} = q_{t+1}^*(n-1) - q_{t+1}(n-1) - (q_t^*(n) - q_t(n)) - x_t \tag{6}$$

We define the conditional rational expectations for the above returns, or relative bond and currency premia, as follows:

$$\Theta_t^{FX} \equiv \mathbb{E}_t[r_{t+1}^{FX}] = x_t + \mathbb{E}_t[s_{t+1}] - s_t \tag{7}$$

$$\Theta_t^{FX}(n) \equiv \mathbb{E}_t[r_{t+1}^{FX}(n)] = \mathbb{E}_t[q_{t+1}^*(n-1) - q_{t+1}(n-1)] - (q_t^*(n) - q_t(n)) + \mathbb{E}_t[s_{t+1}] - s_t \tag{8}$$

$$\Theta_t^{B} \equiv \mathbb{E}_t[r_{t+1}^{B}(n)] = \mathbb{E}_t[q_{t+1}^*(n-1) - q_{t+1}(n-1)] - (q_t^*(n) - q_t(n)) - x_t \tag{9}$$

Similarly we can define the subjective relative currency and bond premia as

$$\zeta_t^{FX} \equiv \mathbb{E}^S[r_{t+1}^{FX}] = x_t + \mathbb{E}_t^S[s_{t+1}] - s_t \tag{10}$$

$$\zeta_t^{FX}(n) \equiv \mathbb{E}^S[r_{t+1}^{FX}(n)] = \mathbb{E}_t^S[q_{t+1}^*(n-1) - q_{t+1}(n-1)] - (q_t^*(n) - q_t(n)) + \mathbb{E}_t^S[s_{t+1}] - s_t \tag{11}$$

$$\zeta_t^B(n) \equiv \zeta_t^{FX}(n) - \zeta_t^{FX} = \mathbb{E}^S[r_{t+1}^B(n)] = \mathbb{E}_t^S[q_{t+1}^*(n-1) - q_{t+1}(n-1)] - (q_t^*(n) - q_t(n)) - x_t \tag{12}$$

These represents the subjectively expected returns from the above trading strategies. In a standard model, these would represent compensation for risk. However, more broadly they can also include "convenience yields" (see e.g. Jiang et al. 2018) necessary to explain violations from covered interest parity type no-arbitrage conditions (see e.g. Du et al. 2018).

For simplicity we assume all components of the above three equations, relative premia, short rate differentials, relative bond price changes[4] and exchange rate changes (but not the level of the FX rate) are stationary under the subjective measure. For notational convenience we assume these have been demeaned with their unconditional mean. We can iterate the first equation to provide an expression for the level of the FX rate

$$s_t = \mathbb{E}_t^S \sum_{j=0}^{\infty} x_{t+j} - \mathbb{E}_t^S \sum_{j=0}^{\infty} \zeta_{t+j}^{FX} + \lim_{j \to \infty} \mathbb{E}_t^S[s_{t+j}].$$

This states that the level of FX rate reflects the subjectively expected path of short rate differentials and risk premia as well as a permanent component of the FX rate. A similar decomposition under the objective measure has been considered e.g. by Engel (2014) and Jiang et al. (2018); for an early application for the real exchange rate see Clarida and Gali (1994). Note that $\lim_{j \to \infty} \mathbb{E}_t^S[s_{t+j}]$ is generally time-varying and hence s_t is non-stationary. In particular it may feature a unit root.

We can solve an analoguous expression for the relative bond price:

$$q_t^*(n) - q_t(n) = -\mathbb{E}_t^S \sum_{j=0}^{n-1} x_{t+j} - \mathbb{E}_t^S \sum_{j=0}^{n-1} \zeta_{t+j}^B.$$

[4]Note that bond prices not only bond price changes might be stationary though this is not required.

The permanent component does not appear in this expression due to finite maturity. Note that, holding other terms constant, the expected path of short rate differentials affects both the level of FX rate and the relative value of bonds. However, the value of the foreign currency is increasing in expected (foreign - home) short rate differentials but the relative value of the foreign bond is decreasing in expected short rate differentials. This fundamental property will be important for the later results. Plugging these expressions to the formulas for the rational expectations of currency returns we obtain

$$\begin{aligned}
\underbrace{\Theta_t^{FX}}_{\text{Currency premium}} = \underbrace{\zeta_t^{FX}}_{\text{Risk premium differential}} &+ \underbrace{\mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} x_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{\infty} x_{t+1+j}\right]}_{\text{Interest rate misperception effect}} \\
&- \underbrace{\mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} \zeta_{t+1+j}^{FX} - \mathbb{E}_t^S \sum_{j=0}^{\infty} \zeta_{t+1+j}^{FX}\right]}_{\text{Risk premium misperception effect}} + \underbrace{\mathbb{E}_t[\underset{j\to\infty}{lim} \mathbb{E}_{t+1}^S[s_{t+j}] - \underset{j\to\infty}{lim} \mathbb{E}_t^S[s_{t+j}]]}_{\text{Permanent component misperception effect}} \\
&\equiv \zeta_t^{FX} + \Theta_t^{IRM} + \Theta_t^{RPM} + \Theta_t^{PCM},
\end{aligned}$$

where the second equality is simply naming a letter for each component. Similarly for bonds we have

$$\begin{aligned}
\underbrace{\Theta_t^{B}(n)}_{\text{Bond premium differential}} = \underbrace{\zeta_t^{B}(n)}_{\text{Risk premium differential}} &+ \underbrace{\mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{n-2} x_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{n-2} x_{t+1+j}\right]}_{\text{Interest rate misperception effect}} \\
&- \underbrace{\mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{n-2} \zeta_{t+1+j}^{B}(n-j-1) - \mathbb{E}_t^S \sum_{j=0}^{n-2} \zeta_{t+1+j}^{B}(n-j-1)\right]}_{\text{Risk premium misperception effect}} \\
&\equiv \zeta_t^{B}(n) + \Theta_t^{B,IRM}(n) + \Theta_t^{B,RPM}(n).
\end{aligned}$$

A similar decomposition is obtained for $\Theta_t^{FX}(n)$. Now consider, the standard simple linear return forecasting model

$$r_{t+1} = \alpha + \beta f_t + \epsilon_{t+1},$$

where f_t is a forecasting factor, ϵ_{t+1} is a zero mean error term and in place of r_{t+1} we can have either r^{FX}_{t+1}, $r^{FX}_{t+1}(n)$ or $r^{B}_{t+1}(n)$. In any sample T the OLS estimate of β is given by

$$\beta = \frac{Cov(r_{t+1}, f_t)}{Var(f_t)} = \frac{Cov(\Theta_t, f_t)}{Var(f_t)},$$

where $\Theta_t = \mathbb{E}_t[r_{t+1}]$ and the second equality follows from the fact that $r_{t+1} = \Theta_t + r_{t+1} - \Theta_t$, where $r_{t+1} - \Theta_t$ is independent from time t information. The above formula for the OLS estimate of β holds also when the true relationship between r_{t+1} and f_t is not linear. The linearity of covariance and the above decompositions for the bond and currency premia then imply the following decomposition for β

$$\underbrace{\beta}_{\text{Predictability coefficient}} = \underbrace{\beta^{RP}}_{\text{Risk premium differential effect}} + \underbrace{\beta^{IRM}}_{\text{Interest rate misperception effect}} + \underbrace{\beta^{RPM}}_{\text{Risk premium misperception effect}} + \underbrace{\beta^{PCM}}_{\text{Permanent component misperception effect}}.$$

For example $\beta^{FX,IRM}$ is given by

$$\beta^{FX,IRM} = \frac{\mathbb{C}ov(\Theta_t^{FX,IRM}, f_t)}{\mathbb{V}ar(f_t)} = \frac{\mathbb{C}ov(\mathbb{E}_t\left[\mathbb{E}^S_{t+1}\sum_{j=0}^{\infty} x_{t+1+j} - \mathbb{E}^S_t \sum_{j=0}^{\infty} x_{t+1+j}\right], f_t)}{\mathbb{V}ar(f_t)}.$$

Moreover, for the bond premium we naturally have $\beta^{B,PCM}(n) = 0$. This implies that the forecasting power of a factor f_t comes from correlation with the rational risk premium, from correlation with the different misperception parts of the bond or currency premium or both. A similar decomposition can be obtained for a linear model with multiple predicting factors. In this paper we are particularly interested in β^{IRM}, the effect of interest rate misperceptions on bond and currency returns. However, identifying β^{IRM} requires further assumptions about the short rate process.

2.2 Identifying Assumptions

We now describe conditions under which the coefficients described in the previous section can be identified using data on short rates, bond and currency returns and survey expectations on future short rates. We take the short rate process under the subjective and objective measure as exogeneous and later estimate these processes from the data.

The following condition describes the key assumption of the paper:

Condition SE

Under the objective measure the short rate differential x_t follows an AR(p) - process. However, under the subjective measure S, the conditional expectation is given by a sticky expectations process $\mathbb{E}_t^S[x_{t+h}] = k\sum_{j=0}^{\infty}(1-k)^j\mathbb{E}_{t-j}[x_{t+h}]$

As in Coibion and Gorodnichenko (2015), we further focus on the case $p = 1$, that is assume that under the objective measure the short rate difference x_t follows an AR(1) process. This gives a good fit to observed data on short rates but we study the robustness of the results to alternative specifications in the appendix. We can rewrite the sticky expectations process as follows:

$$\mathbb{E}_t^S[x_{t+h}] = k\mathbb{E}_t[x_{t+h}] + (1-k)\mathbb{E}_{t-1}^S[x_{t+h}].$$

As argued by Coibion and Gorodnichenko (2015) this process gives a good fit to survey data on short rates. If beliefs are rational $k = 1$ and the subjective and objective expectations coincide. However, typically $0 < k < 1$ 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.

Note that under the assumption that the objective data is given by an AR(1)-process we have

$$\mathbb{E}_t[x_{t+h}] = \lambda^h x_t.$$

Here $-1 < \lambda < 1$; there is no constant because the variables are demeaned. From the initial definition of a sticky expectations process it then follows:

$$\mathbb{E}_t^S[x_{t+h+j}] = \lambda^j\mathbb{E}_t^S[x_{t+h}]$$

and hence also

$$\mathbb{E}_t^S[x_{t+h}] = k\mathbb{E}_t[x_{t+h}] + (1-k)\lambda\mathbb{E}_{t-1}^S[x_{t+h-1}],$$

this expression is used for deriving some of the following results. How could k be estimated? As in Coibion and Gorodnichenko (2015), it is useful to consider the following regression:

$$x_{t+h} - \mathbb{E}_t^S[x_{t+h}] = \alpha^{FR} + \beta^{FR}[\mathbb{E}_t^S[x_{t+h}] - \mathbb{E}_{t-1}^S[x_{t+h}]] + e_{t+h}. \tag{13}$$

Here we regress the forecast error for the short rate differential on the corresponding forecast revision. As explained in the appendix, the model implies that $\alpha^{FR} = 0$ and $\beta^{FR} = \frac{1-k}{k}$. In a rational model $k = 1$, $\beta^{FR} = 0$ and forecast errors are not predictable. More generally, a positive (negative) coefficient for the regression indicates underreaction (overreaction).

This condition fully pins down the effect of short rate forecast errors on bond and currency returns. In particular we have the following proposition:

Proposition 1 (Condition SE and the Term Structure of Carry Trade Returns)**.** *Assume condition SE holds (under $p = 1$). Now the interested rate misperception part of the FX premia are given by*

$$\Theta_t^{FX,IRM} = \left[1 + \frac{\lambda k}{1-\lambda}\right]\left[\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}]\right] \tag{14}$$

$$\Theta_t^{FX,IRM}(n) = \frac{k\lambda^n}{1-\lambda}\left[\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}]\right]. \tag{15}$$

Similarly the interest rate misperception part for the (population OLS estimate of) slope coefficient in a predictability regression with the short-rate differential, $f_t = x_t$ is

$$\beta^{FX,IRM} = \left[1 + \frac{\lambda k}{1-\lambda}\right]\frac{(1-k)(\lambda-\lambda^3)}{1-(1-k)\lambda^2} > 0. \tag{16}$$

$$\beta^{FX,IRM}(n) = \frac{k\lambda^{n-1}}{1-\lambda}\frac{(1-k)(\lambda-\lambda^3)}{1-(1-k)\lambda^2} > 0. \tag{17}$$

$\beta(n)^{FX}$ decays at rate λ^n and approaches zero as $n \to \infty$. Similarly $\Theta_t^{FX,IRM}(n) \to 0$ as $n \to \infty$. Moreover $\beta^{B,IRM}(n) = \beta^{FX,IRM}(n) - \beta^{FX,IRM} < 0$.

Note that assuming condition SE $\Theta_t^{FX,IRM}(n)$ and $\beta(n)^{FX,IRM}$ tend to zero as $n \to \infty$. That is the term structure of the interest rate component of FX premia is downward sloping. On the other hand, as explained by Lustig et al. (2019), standard models typically do not imply that the rational risk premium component $\zeta_t^{FX}(n)$ is downward sloping. This suggests that allowing for sticky short rate expectations might be key to understanding why the actual term structure of FX premia is downward sloping. In particular if the other components are numerically small this property will hold for the actual FX premia and carry trade returns.

Proposition 1 implies that condition SE alone is sufficient to identify the interest rate misperception component of the slope coefficient in a predictive regression with relative short rate. However, obtaining a full expression for the slope coefficient requires further assumptions. One option is to impose assumptions that effectively set the other components to zero. In particular consider the following condition:

Condition CRP

The risk premia are constant in time under the subjective measure: ζ_t^{FX}, $\zeta_t^{FX}(n)$ and $\zeta_t^{B}(n)$ are constant [5]

In a rational risk based model, all time variation in expected returns would be due to a time-varying risk premium. Condition CRP effectively eliminates this channel. Note that in our general framework it also implies that the risk premium misperception components are zero. This means that all time-variation in returns under objective beliefs is due to misperceptions about future short rates and the permanent component of the FX rate. This latter effect can be muted using the following assumption:

Condition NLRM

The permanent component misperception effect is zero $\mathbb{E}_t[\lim_{j\to\infty} \mathbb{E}_{t+1}^S[s_{t+j}] - \lim_{j\to\infty} \mathbb{E}_t^S[s_{t+j}]] = 0.$

This condition is naturally satistied for example when the investors have correct long-run beliefs. Assuming both condition CRP and NLRM now implies that all time varition in objectively expected returns is due to misperceptions about relative short rates. This implies the following proposition:

[5] Assuming any two implies the third condition as $\zeta_t^B(n) = \zeta_t^{FX}(n) - \zeta_t^{FX}$

Proposition 2 (Condition CRP and the Term Structure of Carry Trade Returns)**.** *Assume conditions SE holds (under $p = 1$). Furthermore assume conditions CRP and NLRM hold. Now the FX premia are given by*

$$\Theta_t^{FX} = \left[1 + \frac{\lambda k}{1-\lambda}\right]\left[\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}]\right] \tag{18}$$

$$\Theta_t^{FX}(n) = \frac{k\lambda^n}{1-\lambda}\left[\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}]\right]. \tag{19}$$

Similarly the (population OLS) estimate of the slope coefficient in a predictability regression with the short-rate differential, $f_t = x_t$ is

$$\beta^{FX} = \left[1 + \frac{\lambda k}{1-\lambda}\right]\frac{(1-k)(\lambda - \lambda^3)}{1-(1-k)\lambda^2} > 0. \tag{20}$$

$$\beta(n)^{FX} = \frac{k\lambda^{n-1}}{1-\lambda}\frac{(1-k)(\lambda - \lambda^3)}{1-(1-k)\lambda^2} > 0. \tag{21}$$

$\beta_1(n)^{FX}$ decays at rate λ^n and approaches zero as $n \to \infty$. Similarly $\Theta_t^{FX}(n) \to 0$ as $n \to \infty$.

The results of this proposition look similar to those of the previous one. However, there is a crucial difference. The results of the previous proposition concern the interest rate misperception component of the currency premia and predictability coefficient. Proposition 2 instead shows that by further imposing conditions CRP and NLRM, the bond and currency risk premia and the related predictability coefficients correspond to these same interest rate misperception components. This is because these conditions imply that the other components are zero. This also implies that the actual term structure of FX premia is downward sloping.

The above conditions are less restrictive than they might initially sound. Note that they do not imply that risk premia are constant under the objective measure, but rather that all time variation in objectively measure risk premia are caused by misperceptions concerning short term interest rates. Hassan and Mano (2017) argue that a full model of the FX premium needs to have both a persistent cross-currency component as well as time-varying part that explains why increases in relative foreign interest rates lead to higher foreign currency returns. The above framework satisfies these requirements. However, note that we only provide a theory about time-variation in bond and currency premia but not about persistent cross-currency differences in these premia.

Moreover, for the predictions to hold especially qualitatively, it is not required that these conditions hold exactly. Rather it is only necessary that time variation in risk premia, misperceptions concerning future risk premia and long-term value of the FX rate are small enough to not offset the effects of short rate misperceptions.

Conditions SE, CRP and NLRM imposed in Proposition 2 can also be used to identify predictability coefficients related to alternative predictors than just the relative level of short rates. In particular they imply that the slope coefficient in a regression with a relative yield spread predictor has the opposite sign than the slope coefficient in a regression with relative short rate. This is because, assuming constant risk premia, the yield spread tends to be low when short rates are high.

To illustrate the logic behind the results, we first show the evolution of the yield curve and exchange rate after a shock that increases the foreign interest rate. Figure 1 plots the impulse responses to an interest rate shock assuming conditions SE, CRP and NLRM[6]. When foreign interest rates increase above home interest rates, forecasters update their relative short rate forecasts upward but not as much as a rational forecaster would do. Because long term interest rates are averages over expected short rates, they increase but less than short rates, so the relative yield curve becomes downward sloping. The price of a long-term bond falls but by less than according to rational expectations. The foreign currency appreciates but by less than predicted by rational expectations. However, in the long-run expectations converge to rational values. During the interim period, a high interest rate predicts positive returns for the foreign currency but low relative returns for long-term foreign bonds.

Given conditions SE, CRP and NLRM, 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 2 shows the decay pattern for relative carry trade returns for different values of the persistence parameter λ[7]. As explained before, the downward sloping term structure emerges because variation in expected bond returns offsets variation in expected currency returns.

[6]The figure assumes the long-run log-exchange rate is 0 so here $s_t = \sum_{i=0}^{\infty} \mathbb{E}_t^S[x_{t+i}]$ and $q_t^*(n) - q_t(n) = -\sum_{i=0}^{n-1} \mathbb{E}_t^S[x_{t+i}]$. The impulse responses are computed using the benchmark calibration derived later.

[7]This shows the relative profitability / predictability coefficient. That is coefficient for the short maturity carry trade is normalized to 1.

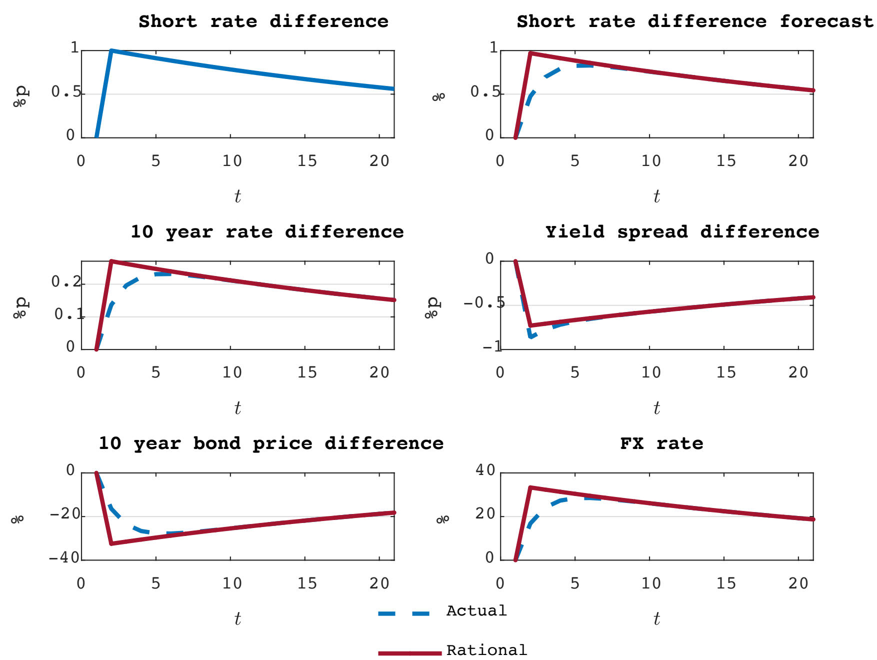

Figure 1 Impulse responses to a shock to the foreign interest rate when short rate forecast errors drive all variation in objective premia. Time is measured in months.

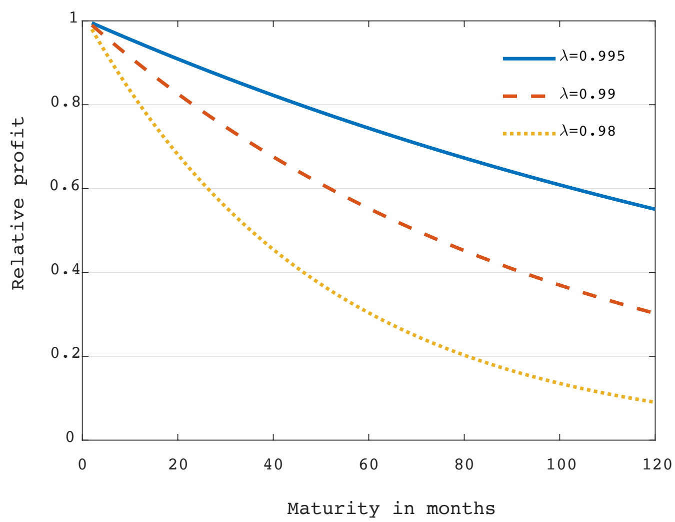

Figure 2 The term-structure of carry trade when short rate forecast errors drive all variation in objective premia.

Figure 3 shows the slope cofficient β^{FX} as a function of both k and λ given conditions SE, CRP and NLRM. The coefficient is positive. For typical parameter values β^{FX} is decreasing in k and increasing in λ. The benchmark calibration used later predicts $\beta_1^{FX} \approx 0.99$

Figure 4 shows the slope cofficient of a regression of relative returns of 10 year bonds on short rate differential x_t assuming conditions SE, CRP and NLRM. This is also given by $\beta^{FX}(n) - \beta^{FX}$. The cofficient is negative. For typical parameter values the slope coefficient is increasing in k and λ. The benchmark calibration discussed later predicts $\beta_1^{FX}(n) - \beta_1^{FX} \approx -0.7$. This opposite predictability in bond returns largely offsets the predictability in currency returns so that there is little predictability in the returns of carry trades implemented with 10 year bonds.

It can be shown that the model predicts the opposite patterns when relative yield spreads are used as predictors. A high slope of the yield curve predicts low currency returns but high bond returns. This occurs because the slope of the yield curve tends to be high when interest rates are low.

Finally, under conditions SE, CRP and NLRM, the model implies that foreign currency returns tend to be particularly high and bond returns low when foreign short rates have recently increased relative to past values. This is formalized in the following proposition:

Proposition 3. *Assume conditions SE, CRP and NLRM hold. Define the average past short rate difference as:* $\bar{x}_t \equiv x_t + (1-k)\lambda x_{t-1} + (1-k)^2\lambda^2 x_{t-2} + \ldots$. *Consider the regressions*

$$r_{t+1}^{FX} = \alpha^{FX} + \beta_1^{FX} x_t + \beta_2^{FX} \bar{x}_{t-1} + \varepsilon_{t+1} \tag{22}$$

and

$$r_{t+1}^{B}(n) \equiv r_{t+1}^{FX}(n) - r_{t+1}^{FX} = \alpha^{B}(n) + \beta_1^{B}(n) x_t + \beta_2^{B}(n) \bar{x}_{t-1} + \varepsilon_{t+1}^{n} \tag{23}$$

The (population OLS) estimate of β_1^{FX} *is positive, of* β_2^{FX} *is negative, of* $\beta_1^{B}(n)$ *is negative and of* $\beta_2^{B}(n)$ *is positive.*

Proof: see appendix.

The forecast wedge $\mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}]$ is particularly wide when x_t is high relative to the past short-rate differences x_{t-1}, x_{t-2}, x_{t-3} and so on. That is, expectational errors concerning short rates are particularly large after recent short rate shocks. On the other hand, given our conditions, the rationally expected currency return is strictly increasing in this forecast wedge and the expected bond return is decreasing. This implies that high

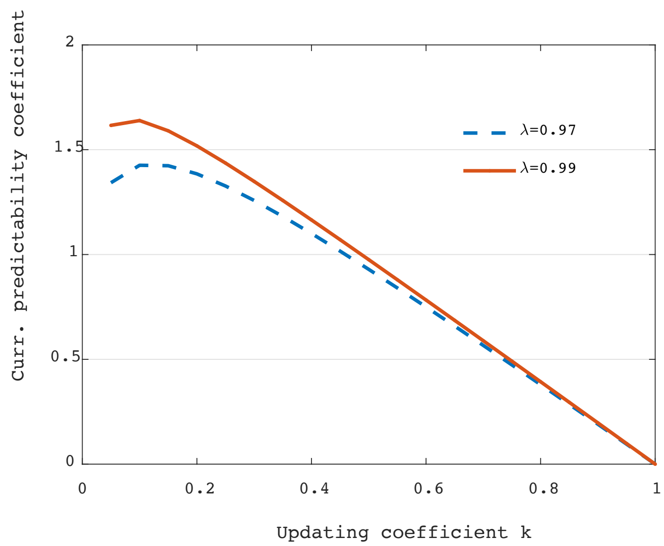

Figure 3 The currency return predictability coefficient as a function of k and λ when short rate forecast errors drive all variation in objective premia.

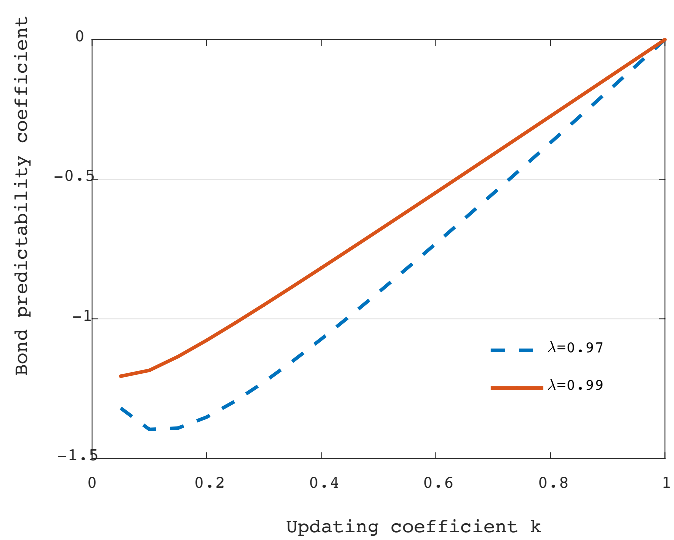

Figure 4 The bond return predictability coefficient as a function of k and λ when short rate forecast errors drive all variation in objective premia.

short-term interest rates relative to past short rates should predict high returns for a currency but low returns for the corresponding long-term bond. This explains why the slope coefficient on the past average short rate difference $\bar{x}_{t-1}$ has the opposite sign than the slope coefficient on the short rate difference x_t. Imposing only condition SE, the above result holds for the interest rate misperception component of bond and currency.

Identifying Expectational Errors

Expectational errors such as the difference between the actual exchange rate and its subjective expectation, $s_{t+j} - \mathbb{E}_t^S[s_{t+j}]$, should be unpredictable under rational expectations. However sticky short rate expectations also imply predictability patterns for expectational errors concerning FX rates and long-term interest rates. However, we need further assumptions to separate the effects of short rate misperceptions on other types of misperceptions. In particular we could impose condition CRP and NLRM that rule out these other effects. However, instead of condition CRP, the following weaker condition is sufficient for most of these results:

Condition NRPM *There are no misperceptions about future values of risk premia Θ_t^{FX} and $\Theta_t^{FX}(n)$.*

That is the risk premium misperception and permanent component misperception effects are zero. Note that condition NRPM is implied by condition CRP. However, it is clearly weaker as it does not require that risk premia are constant under subjective beliefs. This condition is sufficient to yield the following result:

Proposition 4 (Matching Survey Data on Currencies)**.** *Assume conditions SE, NRPM and NLRM hold. Consider the following regression*

$$s_{t+j} - \mathbb{E}_t^S[s_{t+j}] = \alpha + \beta x_t + e_{t+j}. \tag{24}$$

The (population OLS) estimate of β is positive.

Proof: see appendix.

The intuition is illustrated in figure 1. When foreign interest rates increase, the forecasters are sluggish at updating their predictions and the subjective interest rate forecast falls below the rational forecast. Similarly, the FX rate falls below its rational value. The gradual convergence of the

FX rate and interest rate forecasts to rational values leads to an unexpected appreciation pattern in the value of the foreign currency.

Moreover assuming the stricter condition CRP, the correlation between the yield spread and short rate level is negative. For example as can be seen from figure 1, the increase in interest rates leads to a decline in yield curve slope. This implies that forecasters overestimate the future strength of currencies with steep yield curves, so that if we replace short rate in the previous regression with yield spread the predictability coefficient is negative.

Then consider the regression:

$$q_{t+j}(n)^* - \mathbb{E}_t^S[q_{t+j}(n)^*] - (q_{t+j}(n) - \mathbb{E}_t^S[q_{t+j}(n)]) = \alpha + \beta x_t + e_{t+j}. \tag{25}$$

Using similar arguments it can be shown that conditions SE, NRPM and NLRM imply that $\beta < 0$, that is when short-term home interest rates are relatively high, forecasters overpredict the relative future value of foreign bonds. The opposite prediction, $\beta > 0$, is obtained when long-term interest rates are used on the LHS of the equation. Similarly the opposite prection is obtained when conditions SE, CRP and NLRM hold and the slope of the yield curve is used on the RHS of the equation.

3 Empirical Evidence

We now turn to empirically test the model predictions and quantifying the effect of interest rate misperceptions on bond and currency returns.

3.1 Data

We first briefly describe the data used. We focus on the G10 currencies of Australia, Canada, Germany, Japan, Norway, New Zealand, Sweden, Switzerland, U.K. and U.S. We utilize FRED to obtain data on end of month FX rates and interest rates on 3 month and 10 year government securities.

We calibrate the agents' expectations using survey data. Consensus economics provides a monthly report of forecasts for 3 month and 10 year interest rates as well as FX rates. Following Coibion and Gorodnichenko (2012), we average over the forecasts provided by different financial institutions.[8]

[8]This approach is taken in most other papers. It is still unclear whether the results would be different if considering individual forecasts. Bordalo et al. (2019) argue that the

	AUS	CAN	GER	JAP	NOR	NZ	SWE	CH	UK	US
					Bonds Index					
Start	85M1	85M1	85M1	85M1	NA	85M1	85M1	85M1	85M1	85M1
End	19M2	19M2	19M2	19M2	NA	19M2	19M2	19M2	19M2	19M2
Obs	410	410	410	410	NA	410	410	410	410	410
					Consensus					
Start	NA	91M5	91M5	91M5	98M6	NA	94M12	98M6	91M5	89M10
End	NA	19M4	19M4	19M4	19M4	NA	19M4	19M4	19M4	19M4
Obs	NA	336	336	336	251	NA	293	252	336	355

Table 1 Start and end dates for the bond index and survey data.

Forecasts are available for all countries except Australia and New Zealand. Note that the use of professional forecasts might provide a conservative estimate of the biases reflected in asset prices.

We calculate bond returns using Citigroup government bond local currency 10 year indices available for all countries except Norway.[9]

The start and end dates for the bond indices and survey data are given in table 1, where we also report the number of observations.

3.2 Calibration

To quantify the importance of interest rate misperceptions, we need to calibrate the process for short rate differentials x_t. Given condition SE we only need to find the persistence parameter λ and the underreaction coefficient k.

We choose US as the home country and construct a monthly series of the interest rate differential with respect to the other countries (foreign - US rate). We estimate the AR(1) persistence parameter using OLS. We consider the process separately for each country as well as for a panel with all the countries. Note that taking differences removes the common secular

underreaction result is partly driven by averaging though they still find underreaction in individual short rate forecasts. However, averaging might reduce measurement error (see e.g. Juodis and Kucinskas (2019)). Gabaix (2019) argues that agents tend to underreact especially to shocks to persistent processes; short rates are indeed very persistent. For additional discussion see also Bouchaud et al. (2018).

[9]The downside of bond indices is that they are based on coupon bonds, while the theoretical results are for zero-coupon bonds. However, the theoretical predictions hold qualitatively for coupon bonds. Moreover, in unreported robustness checks we obtain similar results using the zero-coupon yield curve data set of Wright (2011), for market data see also the results in Lustig et al. (2019). Moreover, the difference between the predictions for coupon and zero-coupon bonds is small according to simulations. The benefit of using bond indices is that they are free from approximation error in common interpolation procedures and corresponding returns are tradable.

	$\hat{\lambda}_0$	$s.e$	$\hat{\lambda}_1$	$s.e$	R^2
panel			0.987***	0.008	0.979
CAN	0.020	0.023	0.957***	0.016	0.932
GER	-0.016	0.022	0.994***	0.013	0.988
JAP	-0.032	0.017	0.989***	0.008	0.992
NOR	0.014	0.038	0.989***	0.016	0.975
SWE	-0.017	0.027	0.991***	0.011	0.980
CH	-0.014	0.021	0.988***	0.011	0.974
UK	0.005	0.022	0.981***	0.015	0.975

Table 2 shows the results from regressing monthly 3 month yield differential (foreign minus US) on its first lag. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

downward trend in interest rates. The resulting persistence parameters are given in table 2. Interest rate differentials are highly persistent with estimates ranging between 0.96 and 0.99. We choose the panel estimate $\lambda \approx 0.99$ as the baseline calibration.

We then need an estimate of the underreaction coefficient k. For this purpose we regress the forecast error for the short rate differential on the corresponding forecast revision using the 12 month forecasts. As explained before and in greater detail in the appendix, the model implies that the slope coefficient in this regression is $\beta^{FR} = \frac{1-k}{k}$.

Table 3 shows the results from this regression along with the implied values for k. We use the panel estimate $k \approx 0.49$ as the baseline calibration.

With a k above one, indicating overreaction, Canada seems to be an outlier but we still include it in the panel regression. Most of the country specific coefficient values are close to each other. Indeed with the exception of Canada, none of the country -level values are statistically different from the panel estimate.

3.3 Short-rate Misperceptions and Bond and Currency Returns

We start by replicating the four key predictability regressions in Lustig et al. (2019). These include regressions of bond returns either on short-rate differentials or on yield spread differentials. According to Proposition 2 and the discussion in section 2, assuming conditions CRP and NLRM, the slope coefficient in the regressions involving short term rates should be negative and the slope coefficient in the regressions with yield spreads

	$\hat{\beta}_0$	*s.e*	$\hat{\beta}_1$	*s.e*	R^2	*implied k*
panel			1.059**	0.391	0.035	0.486
CAN	0.282	0.164	-0.230	0.217	0.003	1.299
GER	0.304*	0.186	1.628**	0.684	0.074	0.380
JAP	0.331*	0.190	1.771***	0.535	0.083	0.361
NOR	0.351	0.279	1.995***	0.720	0.089	0.334
SWE	-0.319	0.238	1.461***	0.582	0.055	0.406
CH	0.305	0.185	0.972*	0.530	0.028	0.507
UK	0.200	0.171	0.564	0.373	0.008	0.639

Table 3 shows the results from regressing the difference in forecast error (foreign minus US) from forecasting spot 3 month 12 months ahead on the difference in forecast revisions. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

positive. Similarly the same conditions imply that the slope coefficient in the regressions involving yield spreads should be positive. More generally, Proposition 1 implies that the interest rate misperception component of the regression with short rates should be negative.

Table 4, panel A, gives the results for the bond return regressions both for individual countries as well as for the fixed effects panel regressions. The signs of the slope coefficients are as predicted by Proposition 2. The results for the panel regressions are statistically significant. The slope coefficients in the regression with short-rate differentials are similar to those reported by (Lustig et al., 2019) but the slope coefficients in the regressions with yield spreads are somewhat smaller.

We also consider regressing currency returns on short rate differentials and yield spread differentials. According to Proposition 2 and the discussion in section 2, assuming conditions CRP and NLRM, the slope coefficient in the first regression should be positive and the slope coefficient in the second regression negative. More generally, Proposition 1 implies that the interest rate misperception component should be positive in the first regression is positive. The results are given in table 4, panel B. The signs in the panel regressions are statistically significant and as predicted by Proposition 2. The absolute magnitudes of the slope coefficients are somewhat smaller than those in Lustig et al. (2019), especially when using yield spread differentials as predictors.

Table 5 summarizes the results for the panel regressions in Table 4. Here we also show the slope coefficient from a regression of relative bond

PANEL A: Bond Returns										
	3 month rate					yield slope				
	$\hat{\beta}_0$	*s.e*	$\hat{\beta}_1(3)$	*s.e*	R^2	$\hat{\beta}_0$	*s.e*	$\hat{\beta}_1$	*s.e*	R^2
panel			-1.259***	0.253	0.019			1.250**	0.480	0.008
AUS	0.002**	0.001	-1.348***	0.543	0.028	0.001	0.001	2.095**	1.024	0.020
CAN	0.001*	0.000	-1.895***	0.530	0.039	0.001*	0.000	2.554***	0.722	0.034
GER	-0.000	0.000	-0.845**	0.347	0.008	-0.000	0.000	0.598	0.500	0.003
JAP	-0.002***	0.001	-1.212***	0.389	0.089	-0.000	0.000	0.405	0.478	0.001
NZL	0.002**	0.001	-1.588***	0.503	0.016	-0.001	0.001	0.948	0.720	0.001
SWE	0.001	0.000	-1.111**	0.542	0.021	0.001	0.001	1.407	1.038	0.012
CH	-0.002***	0.000	-1.045**	0.513	0.007	-0.000	0.000	1.036	0.666	0.004
UK	0.001*	0.000	-1.427***	0.481	0.021	0.000	0.000	1.072*	0.653	0.004

PANEL B: Currency Returns										
	3 month rate					yield slope				
	$\hat{\beta}_0$	*s.e*	$\hat{\beta}_1(3)$	*s.e*	R^2	$\hat{\beta}_0$	*s.e*	$\hat{\beta}_1$	*s.e*	R^2
panel			1.489**	0.476	0.015			-1.943*	1.019	0.010
AUS	-0.001	0.002	1.477**	0.581	0.014	-0.000	0.002	-2.085	1.388	0.008
CAN	0.002	0.001	1.159*	0.719	0.006	0.002**	0.001	-0.577	1.267	0.001
GER	0.005***	0.002	-.455	1.178	0.001	0.005**	0.002	0.196	1.449	0.000
JAP	0.005**	0.002	2.972***	0.933	0.030	-0.004**	0.001	-4.122***	1.153	0.029
NZL	-0.003	0.002	2.023***	0.419	0.052	-0.001	0.002	-3.467***	1.099	0.050
SWE	-0.002	0.002	0.849	0.897	0.004	-0.001	0.001	0.479	1.595	0.001
CH	0.003*	0.002	1.754*	1.046	0.008	-0.000	0.001	-2.531*	1.418	0.012
UK	-0.000	0.002	1.517*	0.944	0.010	-0.000	0.001	-1.963	1.640	0.007

Table 4 shows the results from regressing the relative bond (foreign minus US) and currency returns on short-rate and yield spread differences. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

Coefficient	Data	β^{IRM} under SE	Model SE+CRP+NLRM
β, LHS: Currret, RHS: x_t	1.489	0.99 (66 %)	0.99
β, LHS: Currret, RHS: yield spread	-1.943	NA	-3.3
β, LHS: $r^{FX}_{t+1}(120)$, RHS: x_t	0.23	0.29 (126 %)	0.29
β, LHS: $r^{FX}_{t+1}(120)$, RHS: yield spread	-0.69	NA	-0.98
β, LHS: Bondret, RHS: x_t	-1.259	-0.7 (56 %)	-0.7
β, LHS: Bondret, RHS: yield spread	1.250	NA	2.339
Volatility ratio, 10 year rate, 3 month rate	0.67	NA	0.57

Table 5 shows key statistics measured from the data (panel regressions) as well as those predicted by the model.

returns, expressed in the same currency, on short-rate and yield spread differentials. This is mechanically the sum of the slope coefficients in the bond and currency regressions. The table also shows the theoretically implied coefficients using the calibration described earlier.

Overall, the empirical sample values are close to those predicted by the model under sticky short rate expectations and conditions CRP and NLRM. Only the model implied coefficients in the panel regressions with yield spreads are somewhat larger in absolute value than in the data. On the other hand, these estimates are somewhat noisy and for example the coefficients in the spread regressions are larger than those obtained by Lustig et al. (2019) using a sample period starting earlier.

As explained in the theoretical part, only the assumption of sticky short rate expectations is required to identify the interest rate misperception components of the slope coefficients in the regressions with relative short rates. Therefore given our calibration for the short rate process, but otherwise quite generally, short rate misperceptions account for 66% of FX return and 56% of relative bond return predictability related to variation in short rates. Moreover, they explain essentially all of the time-series predictability of carry trades implemented with long maturity bonds, but as explained these returns are small on average.

This contribution of short rate forecast errors can be further explained in the following way. For example, the sample value for the predictability coefficient related to regressing currency returns on short rate differences is 1.489. Given our calibration for the short rate process, the interest rate misperception component is 0.99. Our previous general decomposition then implies that the three other components related to a risk premium, future risk premium misperceptions and misperceptions about the long-run exchange rate must sum to $1.489 - 0.99 \approx 0.5$. Without further assumptions we cannot separate the relative contribution of these other parts. However, if we impose conditions NMRP and NLRM, that is there are no misperceptions about risk premia or the long-run exchange rate, this 0.5 represents a rational risk premium. Therefore either the rational risk premium is fairly small or it is large but misperceptions about future risk premia and the long-run FX rate offset most of its effects on return predictability.

These results suggest that forecast errors concerning short rates can go a surprisingly long way in explaining the above predictability patterns. As explained by Lustig et al. (2019) standard rational expectations macrofinance models have trouble replicating these patterns. Moreover, these models lack a channel that seems to quantitatively explain most of the predictability of bond and currency returns.

The table also shows the ratio between the volatilities of 10 year rate differentials and 3 month rate differentials. The number is 0.67 in the data as compared to 0.57 predicted by our calibration and assumptions CRP and NLRM. Shiller (1979) explains that empirically the volatility of long-term rates is higher than would be justified by the path of rationally expected short-rates. However, sticky expectations concerning short-rates increase the volatility of long-rates so that the model predicted value is not far from that in the data.[10]

Bond and Currency Returns and Past Short Rates In a sticky expectations model, short rate forecast errors tend to be particularly high after recent short rate changes. This is because it takes time for forecasters to update their predictions. This implies that high short-term interest rates relative to past short rates should predict high returns for a currency but low returns for the corresponding long-term bond, as explained in Proposition 3.

We now test this implication of the sticky expectations model. We construct the past average short rate difference $\bar{x}_t$ using our estimates of k and λ[11]. We then regress relative bond and currency returns on x_t and $\bar{x}_{t-1}$. The results are given in table 6. The slope coefficients on short rate differences are as before though larger in magnitude. However, as predicted by the model the slope coefficient on the average past short rate is negative in the currency regression but positive in the bond regression. Foreign currency returns tend to be particularly high when the foreign short rate has recently increased. Similarly foreign bond returns tend to be particularly high when

[10] Gourinchas and Tornell (2004) also show that sticky short rate expectations can account for the related currency persistence and volatility puzzles (e.g. Backus et al. (1993), Moore and Roche (2002)).

[11] Because we weight the past rates with our estimates of k and λ, this regression is generally subject to a generated regressor problem. However, alternative weighting schemes that do not depend on these estimates yield similar results.

PANEL A: Bond Returns

	$\hat{\beta}_0$	s.e	$\hat{\beta}_1\ (x_t)$	s.e	$\hat{\beta}_2\ (\bar{x}_{t-1})$	s.e	R^2
panel			-1.322***	1.172	3.917***	0.564	0.044
AUS	0.000	0.001	-8.962***	2.118	4.061***	1.038	0.045
CAN	0.000	0.001	-8.228**	3.678	3.560**	1.773	0.052
GER	-0.001	0.001	-9.132***	2.451	4.232***	1.185	0.038
JAP	-0.002**	0.001	-10.880***	3.171	4.975***	1.562	0.022
NZL	0.000	0.001	-7.513***	2.822	3.263**	1.401	0.032
SWE	0.000	0.001	-7.419***	1.373	3.369***	0.722	0.064
CH	-0.002***	0.001	-10.698***	3.359	4.879***	1.589	0.046
UK	-0.001***	0.001	-8.783***	2.339	4.017***	1.164	0.036

PANEL B: Currency returns

	$\hat{\beta}_0$	s.e	$\hat{\beta}_1\ (x_t)$	s.e	$\hat{\beta}_2\ (\bar{x}_{t-1})$	s.e	R^2
panel			4.961**	1.717	-1.838*	0.827	0.015
AUS	-0.001	0.002	7.265***	2.813	-2.797**	1.322	0.045
CAN	0.001	0.001	0.544	3.675	0.397	1.734	0.052
GER	0.005**	0.002	2.201	5.064	-1.233	2.436	0.038
JAP	0.004*	0.002	3.626	6.349	-0.417	3.159	0.022
NZL	-0.001	0.002	6.894**	2.899	-2.582*	1.481	0.032
SWE	-0.001	0.002	3.810*	2.287	-1.796	1.312	0.064
CH	0.003	0.002	7.355	5.604	-2.645	2.802	0.046
UK	0.000	0.002	2.423	3.574	-0.717	1.676	0.036

Table 6 shows the results from regressing the relative bond (foreign minus US) and currency returns on short rate differences and an average of past short rate differences. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

the foreign short rate has recently decreased.

Strictly speaking these predictions require conditions CRP and NLRM. However, they hold more generally for the interest rate misperception components of the slope coefficients. Moreover, they hold for the actual slope coefficients if there is no mechanism large enough that would offset these effects.

These results further support the model mechanism depicted in figure 1 and the idea that bond and currency return predictability patterns are largely *drift* patterns. Here a positive shock to the foreign short rate leads to a slow appreciation of the foreign currency and a sluggish fall in the value of foreign bond.

The above findings are related to the delayed overshooting (Eichenbaum and Evans, 1995) and FOMC post announcement drift (see e.g. Brooks et al. (2019)) patterns documented in the previous literature. Delayed overshooting refers to the fact that the response of the FX rate to interest rate shocks is hump-shaped. A contractionary shock to US monetary policy induces a gradual appreciation of the US dollar followed by depreciation. The model with conditions SE, CRP and NLRM implies an identical exchange rate process to that in Gourinchas and Tornell (2004). They show that this process can account for the delayed overshooting puzzle.

A pattern similar to the delayed overshooting puzzle of currencies is the FOMC post announcement drift in bond markets. Here the yields of long-maturity treasuries respond sluggishly to changes in the Federal Funds Rate. In concurrent work, Brooks et al. (2019) argue that slow updating concerning short-term interest rates can explain the FOMC post announcement drift. This pattern is also generated by our model as can be seen in the FX impulse response plotted in figure 1.

3.4 Predicting Expectational Errors in Survey-based Expectations

Expectational errors should be unpredictable under rational expectations. However, a key prediction of our sticky expectations model is that the same variables that predict bond and currency returns also predict expectational errors concerning FX rates and long-term interest rates. These predictions do not require that subjective risk premia are constant, only that there are no other misperceptions large enough to offset the short rate misperceptions channel.

Bacchetta et al. (2009) find support for the model prediction that short-rate differentials predict expectational errors concerning FX rates. They also find that yield spreads predict expectational errors concerning long-term interest rates but do not offer an explanation for these findings.

We verify these predictions using an alternative dataset and a different sample period. Moreover, we find support for the additional prediction that yield spread differentials explain expectational errors concerning FX rates.

The results are given in table 7. As predicted by the model a high short-rate differential between the foreign and home country predicts that the foreign currency will turn stronger than expected. The opposite prediction is obtained when using yield spread differentials as the explanatory variable. The results are given in table 7. The results for panel regressions are statistically significant and as predicted by the model.

The model also predicts that when short-rates are high, long-term interest rates (bond prices) will turn higher (lower) than predicted. Moreover, according to the sticky expectations model a high yield spread predicts that long-term interest rates will turn lower than expected. The results for this regression are given in table 8. Using a panel regression, we find positive

Forecast Errors FX: 12 Months Ahead										
	3 month rate differential					term spread differential				
	$\hat{\beta}_0$	$s.e$	$\hat{\beta}_1(3)$	$s.e$	R^2	$\hat{\beta}_0$	$s.e$	$\hat{\beta}_1$	$s.e$	R^2
panel			0.012*	0.006	0.040			-0.016*	0.008	0.037
CAD	0.021*	0.012	0.003	0.006	0.003	0.020*	0.011	-0.003	0.008	0.002
CHF	-0.047***	0.013	0.016***	0.006	0.133	-0.013*	0.008	-0.023***	0.008	0.206
EUR	0.001	0.016	0.023**	0.012	0.094	-0.004	0.017	-0.035**	0.018	0.088
GBP	0.012	0.015	-0.004	0.008	0.006	0.022*	0.013	0.013	0.010	0.030
JPY	-0.051***	0.019	0.024***	0.007	0.171	0.029*	0.017	-0.043***	0.009	0.238
NOK	0.131***	0.045	0.033**	0.016	0.120	0.127***	0.044	-0.040**	0.019	0.107
SEK	0.029***	0.009	-0.004	0.008	0.014	0.026**	0.009	0.014	0.009	0.081

Table 7 shows the results from regressing the difference in forecast error (foreign minus US) from forecasting 3 month spot FX rates 12 months ahead on short-rate and yield spread differentials. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

support for the latter prediction. The slope coefficient in the panel regression with short rate is insignificant. However, it becomes significant if we exclude Japan from the sample. Here, the fact that Japan has experienced close to zero interest rates for most of the sample might complicate the relationship between short-rates and expectational errors. Moreover, evidence that short-rates predict expectational errors in long-term rates is provided by Bacchetta et al. (2009).

4 An Affine Term Structure Model

To emphasize generality, the previous sections said little about how risk premia might be determined under the subjective measure. To illustrate our results further and how they map into those of the previous liteture we now make an assumption about the form of the relevant stochastic discount factors. We also make a particular assumption about the belief process that naturally gives rise to sticky short rate expectations.

We now assume that markets are complete. We also consider the case of symmetric countries. This is useful for illustrative purposes and because we focus on time series predictability, not explaining persistent cross-country differences between returns. This model is similar to that presented by Gourinchas and Tornell (2004).

Under the subjective measure S, the home and foreign nominal stochastic

Forecast Errors 10Y Bond: 12 Months Ahead										
	3 month rate					yield slope				
	$\hat{\beta}_0$	$s.e$	$\hat{\beta}_1(3)$	$s.e$	R^2	$\hat{\beta}_0$	$s.e$	$\hat{\beta}_1$	$s.e$	R^2
panel			0.010	0.017	0.019			-0.166**	0.053	0.049
CAN	-0.706***	0.097	0.031	0.026	0.004	-0.494***	0.135	-0.083	0.105	0.009
GER	-0.800***	0.072	0.046**	0.020	0.022	-0.478***	0.137	-0.193**	0.081	0.080
JAP	-0.367***	0.036	-0.124***	0.023	0.109	-0.213*	0.129	-0.240**	0.113	0.102
NOR	-0.798***	0.123	0.070**	0.034	0.036	-0.473***	0.127	-0.255***	0.094	0.136
SWE	-0.713***	0.080	-0.072***	0.024	0.035	-0.385*	0.221	-0.352***	0.122	0.107
CH	-0.554***	0.058	0.059*	0.033	0.011	-0.357***	0.133	-0.136	0.117	0.019
UK	-0.502***	0.094	-0.005	0.019	0.000	-0.486***	0.111	-0.054	0.075	0.005
USA	-0.754***	0.084	0.040*	.023	0.001	-0.297	0.192	-0.186**	0.090	0.058

Table 8 shows the results from regressing the difference in forecast error from forecasting 10 year interest rates 12 months ahead on the short-rates and yield slopes. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

discount factors (SDFs), $M_{t,t+1}$ and $M^*_{t,t+1}$, follow symmetric (conditionally) log-normal processes

$$log(M_{t,t+1}) \equiv m_{t,t+1} = -logR - \frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^2}{2} - \bar{z}_t - \frac{\sigma_\epsilon^2 \varphi_t^2}{2} - z_t - \bar{\varphi}_t \bar{\epsilon}_{t+1} - \varphi_t \epsilon_{t+1} \quad (26)$$

$$log(M^*_{t,t+1}) \equiv m^*_{t,t+1} = -logR - \frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^{*2}}{2} - \bar{z}_t - \frac{\sigma_\epsilon^2 \varphi_t^{*2}}{2} - z^*_t - \bar{\varphi}^*_t \bar{\epsilon}_{t+1} - \varphi^*_t \epsilon^*_{t+1}. \quad (27)$$

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[12] 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. These states can represent either deep structural state variables or reduced form factors often used in term structure models.

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

$$\mathbf{z}_t = \Lambda \mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \quad (28)$$

where

[12]Note that we assume countries are symmetric and the shocks ϵ_t and ϵ^*_t have the same variance σ_ϵ^2.

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

Here $0 < \lambda < 1$ and $0 < \bar{\lambda} < 1$. The home short rate is simply $logR + z_t + \bar{z}_t$ and the foreign rate $logR + z_t^* + \bar{z}_t$. This implies that under the objective measure the short rate differential x_t evolves

$$x_t = \lambda x_{t-1} + \tilde{\epsilon}_t,$$

where $\tilde{\epsilon}_t \equiv \epsilon_t - \epsilon_t^*$. Hence as in condition SE, the actual short rate differential process follows an AR(1) process. More spefically, our distributional assumptions imply that it follows a Gaussian AR(1)-process.

On the other hand, the investors believe that the state variables follow (i.e. their S-dynamics are given by):

$$\mathbf{z}_t = \mathbf{l}_t + \boldsymbol{v}_t, \tag{29}$$

$$\mathbf{l}_t = \Lambda \mathbf{l}_{t-1} + \boldsymbol{\epsilon}_t. \tag{30}$$

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.[13] Note that the agents correctly observe all the state variables but misperceive their law of motion. In particular they erroneously believe that the effects 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

$$\varphi_t = \varphi_0 + \varphi_1 z_t + \varphi_2 \bar{z}_t \quad \bar{\varphi}_t = \bar{\varphi}_0 + \bar{\varphi}_1 z_t + \bar{\varphi}_2 \bar{z}_t$$

$$\varphi_t^* = \varphi_0^* + \varphi_1^* z_t + \varphi_2^* \bar{z}_t \quad \bar{\varphi}_t^* = \bar{\varphi}_0^* + \bar{\varphi}_1^* z_t + \bar{\varphi}_2^* \bar{z}_t$$

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 5 (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*

[13] All parameters in the model are assumed to be known by the agents.

$$\mathbb{E}_t^S[\boldsymbol{z}_{t+1}] = \begin{bmatrix} \lambda(1-k_t) & 0 & 0 \\ 0 & \lambda(1-k_t) & 0 \\ 0 & 0 & \bar{\lambda}(1-\bar{k}_t) \end{bmatrix} \mathbb{E}_{t-1}^S[\boldsymbol{z}_t] + \begin{bmatrix} \lambda k_t & 0 & 0 \\ 0 & \lambda k_t & 0 \\ 0 & 0 & \bar{\lambda}\bar{k}_t \end{bmatrix} \boldsymbol{z}_t, \tag{31}$$

$$\mathbb{E}_t^S[x_{t+1}] = \lambda(1-\tilde{k}_t)\mathbb{E}_{t-1}^S[x_t] + \lambda\tilde{k}_t x_t. \tag{32}$$

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 for simplicity we assume that the estimators have converged to their steady-state values. This assumption is quite standard, see e.g. Gourinchas and Tornell (2004). Note that k_t and $\tilde{k}_t$ are generally different but converge to the same value.

The proposition therefore implies that the subjective expectations for the short rate differential follow a sticky expectations process as in condition SE. We conclude that condition SE holds in the affine model of this section. This is related to the argument in Coibion and Gorodnichenko (2012) that a noisy information model implies sticky expectations.

The bond prices and FX rate can be solved using the following standard pricing conditions:

$$e^{q_t(n)} = \mathbb{E}_t^S[M_{t+1}e^{q_{t+1}(n-1)}],$$

$$e^{q_t(n)^*} = \mathbb{E}_t^S[M_{t+1}^* e^{q_{t+1}(n-1)^*}],$$

$$m_{t+1} + s_{t+1} - s_t = m_{t+1}^*,$$

where the currency pricing equation follows from complete markets. This specification for the SDFs then implies a closed form expression for the bond prices as argued by the following proposition:

Proposition 6 (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*

$$q_t(n) = A(n) + \boldsymbol{B}(n)'\boldsymbol{Y}_t, \tag{33}$$

where $A(n)$ and $\boldsymbol{B}(n)$ are given in the appendix. The foreign prices of zero coupon bonds take an analogous form but the state variables are $\boldsymbol{Y}_t^* = [z_t^*, \bar{z}_t, \mathbb{E}_t^S[z_{t+1}^*], \mathbb{E}_t^S[\bar{z}_{t+1}]]'$.

Proof: see appendix.

As argued in the appendix the currency premium depends on these same state variables. Hence we can view our specification as a six factor affine model with non-standard factor dynamics and special restrictions between the three "true" state variables and their subjective expectations.

To obtain the key results of the previous section, it is necessary to assume only that agents have incorrect beliefs about the common shock. That is we can have $\bar{k} = 1$. However more generally, our affine specification nests a fully rational model as a special case: $\bar{k} = k = 1$. This case occurs when $\sigma_v^2 = \bar{\sigma}_v^2 = 0$ so that the subjective and objective measures coincide.

Condition CRP Condition CRP that is constant risk premia under the subjective measure is obtained as a special case of the above model. This occurs when the time-varying parts of market prices of risk are zero:

$$\varphi_1 = \varphi_2 = \bar{\varphi}_1 = \bar{\varphi}_2 = 0.$$

The condition does not require that the constant parts $\varphi_0 = \bar{\varphi}_0$ are zero. These parameters can still be set to e.g. replicate the average positive slope of the yield curve. Note that with these two parameters we can e.g. match the mean of 10 year rate and short rate exactly. However, in general we cannot match the whole average yield curve perfectly.

Given this special restriction Propositions 2,3 and 4 hold exactly. Note that the affine model of this section naturally satisfies condition NLRM. This is because given a shock to state variables the agents beliefs eventually converge to rational ones.

Rational vs Sticky Expectations Model: An Estimation Exercise To furher demonstrate that accounting for sticky short rate expectations helps in matching the data, we now estimate the above models. We estimate the market price of risk parameters φ_0, φ_1, φ_2 and $\bar{\varphi}_0$, $\bar{\varphi}_1$, $\bar{\varphi}_2$ but calibrate all other parameters. As in the previous section, the persistence parameters λ and $\bar{\lambda}$ can be estimated directly using short rate data and we set $\lambda = \bar{\lambda} = 0.99$. We consider both a rational calibration with $k = \bar{k} = 1$ and sticky expectations calibration with $k = \bar{k} = 0.49$[14].

[14]We set these two parameters to be equal because forecast revisions predict short rate revisions and short rate differential revisions in roughly the same way.

We estimate the market price of risk parameters as follows. We target 4 of the 6 slope parameters in table 5[15]. In particular we consider the regressions with relative bond returns and relative FX returns. We also target the volatility ratio between 10 year rates and 3 month rates. The appendix gives closed-form expression for the model implied coefficients. We use numerical optimization to find the parameters that minize the equally weighted sum of squared deviations between the model implied coefficients and those in the data.

The rational model yields a sum of squared deviations of 0.42. The sum of the absolute values of the four parameters that determine the time-variation in market prices of risk is 5.45. In this case the rational model does fairly well largely because there are six free parameters but five target values.

We then consider the sticky expectations calibration but estimate the same market prices of risk parameters. The model yields a sum of squared deviations of 0.27. Hence the pricing errors of the model fall by 36%. The sum of the absolute values of the four parameters that determine the time-variation in market prices of risk is smaller at 3.00. This is because the sticky expectations model attributes a smaller part of time variation in objective premia to time-variation in market price of risk. We conclude that accounting for sticky expectations in a standard affine term structure model helps in matching data on bond and currency returns.

Note that the sticky expectations version of the model is also broadly more consistent with the data. In particular it explains why forecast errors about interest rates and currencies are predictable and why short rate changes rather than short rates seem to predict bond and currency returns.

On Consumption Based Asset Pricing Models The above affine model does not give a direct economic interpretation about the sources of bond or currency risk. However, it nests several consumption based specifications. To see this assume each contry is populated by a representative agent with CRRA prefrences $\beta\frac{C_t^{1-\gamma}}{1-\gamma}$. The real SDF is given by

$$m_{t+1} = log\beta - \gamma\Delta c_{t+1}$$

Consider an endowment economy where log-consumption follows:

$$\Delta c_{t+1} = -z_t - \bar{z}_t + \epsilon_{t+1}$$

[15]Note that these slopes fully determine the remaining two.

and the factors have the law of motion specified as follows. This would be a simple example of a model that is of our affine form and satisfies condition CRP.[16]

Note that the above example abstracts away from inflation. We could distinguish 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)).[17] Empirically shocks to expected inflation contribute much less to the variation in nominal yields than would be predicted by many structural models (Duffee (2018) and Haubrich et al. (2012)).[18]

5 Conclusion

We show that well-documented sluggish updating concerning short rates creates joint predictability patterns in bond and currency markets. These predictability patterns explain most of the variation in expected bond and currency returns driven by variation in short rates and yield spreads.

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.

The analysis bears important policy implications. Monetary policy that affects short rates transmits to bond yields and FX rates at a lag. Including sticky expectations to standard term structure models allows them to better capture the predictability patterns in the data.

References

Bacchetta, P., Mertens, E., and Wincoop, E. V. (2009). Predictability in financial markets: What do survey expectations tell us? *Journal of*

[16] About the relationship between affine term structure models and consumption based models see e.g. Creal and Wu (2020).

[17] 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).

[18] We would therefore expect our state variables to have higher correlations with real variables rather than inflation rates. However, our approach allows for different interpretations concerning these variables.

International Money and Finance, 28:406–426.

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.

Bordalo, P., Gennaioli, N., Ma, Y., and Shleifer, A. (2019). Over-reaction in macroeconomic expectations. NBER Working Paper No. 24932.

Bouchaud, J.-P., Kruger, A., Landier, A., and Thesmar, D. (2018). Sticky expectations and the profitability anomaly. *Journal of Finance*, 74(2):639–674.

Brooks, J., Katz, M., and Lustig, H. (2019). Post-FOMC announcement drift in U.S. bond markets. Working Paper.

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.

Cieslak, A. (2017). Short-rate expectations and unexpected returns in treasury bonds. *Review of Financial Studies*, 31(9):3265–3306.

Clarida, R. and Gali, J. (1994). Sources of real exchange-rate fluctuations: How important are nominal shocks? In *Carnegie-Rochester conference series on public policy*, volume 41, pages 1–56.

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

Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities. *Journal of Political Economy*, 120(1):116–159.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8).

Creal, D. and Wu, C. (2020). Bond risk premia in consumption-based models. *Quantitative Economics, Forthcoming*.

D'Acunto, F., Hoang, D., Paloviita, M., and Weber, M. (2019). Cognitive abilities and expectations. *American Economic Review Papers and Proceedings*, 109:1–5.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, 80:549–560.

Du, W., Tepper, A., and Verdelhan, A. (2018). Deviations from covered interest rate parity. *The Journal of Finance*, 73(3):915–957.

Duffee, G. (2018). Expected inflation and other determinants of treasury yields. *Journal of Finance*, 73(5):2139–2180.

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406.

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. (2014). Exchange rates and interest parity. In *Handbook of international economics*, volume 4, pages 453–522. Elsevier.

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

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.

Gabaix, X. (2019). Chapter 4 - behavioral inattention. In Bernheim, D., DellaVigna, S., and Laibson, D., editors, *Handbook of Behavioral Economics: Applications and Foundations*, volume 2.

Gabaix, X. and Maggiori, M. (2015). International liquidity and exchange rate dynamics. *Quarterly Journal of Economics*, 130(3):1369–1420.

Giacoletti, M., Laursen, K., and Singleton, K. (2018). Learning 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., Hanson, S., Stein, J., and Sunderam, A. (2019). A quantity-driven theory of term premiums and exchange rates. Working paper.

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.

Hassan, T. and Mano, R. (2017). Forward and spot exchange rates in a multi-currency world. *Quarterly Journal of Economics*, 134(1):397–450.

Haubrich, J., Pennacchi, G., and Ritchken, P. (2012). Inflation expectations, real rates, and risk premia: Evidence from inflation swaps. *The Review of Financial Studies*, 25(5):1588–1629.

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.

Jiang, Z., Krishnamurthy, A., and Lustig, H. (2018). Foreign safe asset demand and the dollar exchange rate. NBER Working Paper No. w24439.

Juodis, A. and Kucinskas, S. (2019). Quantifying noise. Working Paper.

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

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. (2019). The term structure of currency carry trade risk premia. *American Economic Review*, 109(12).

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.

Sarno, L., Schneider, P., and Wagner, C. (2012). Properties of foreign exchange risk premiums. *Journal of Financial Economics*, 105(2):279–310.

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

Verdelhan, A. (2010). A habit-based explanation of the exchange risk premium. *Journal of Finance*, 65(1):123–145.

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.

6 Appendix

6.1 Proof of Proposition 1

We have

$$\Theta_t^{FX,IRM} = \mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} x_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{\infty} x_{t+1+j}\right]$$

Note that

$$\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} x_{t+j+1} = x_{t+1} + \frac{1}{1-\lambda}\mathbb{E}_{t+1}^S[x_{t+2}]$$

and

$$\mathbb{E}_t^S \sum_{j=0}^{\infty} x_{t+j+1} = \mathbb{E}_t^S[x_{t+1}] + \frac{1}{1-\lambda}\mathbb{E}_t^S[x_{t+2}]$$

Hence

$$\Theta_t^{FX,IRM} = \mathbb{E}_t\left[x_{t+1} + \frac{1}{1-\lambda}\mathbb{E}_{t+1}^S[x_{t+2}] - \mathbb{E}_t^S[x_{t+1}] - \frac{1}{1-\lambda}\mathbb{E}_t^S[x_{t+2}]\right]$$

$$= \mathbb{E}_t\left[x_{t+1} + \frac{1}{1-\lambda}\mathbb{E}_{t+1}^S[x_{t+2}] - \mathbb{E}_t^S[x_{t+1}] - \frac{\lambda}{1-\lambda}\mathbb{E}_t^S[x_{t+1}]\right]$$

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

.

Similarly

$$\Theta_t^{B,IRM}(n) = -\mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{n-2} x_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{n-2} x_{t+1+j}\right]$$

Note that

$$\mathbb{E}_{t+1}^S \sum_{j=0}^{n-2} x_{t+j+1} = x_{t+1} + \frac{1-\lambda^{n-2}}{1-\lambda}\mathbb{E}_{t+1}^S[x_{t+2}]$$

and

$$\mathbb{E}_t^S \sum_{j=0}^{n-2} x_{t+j+1} = \mathbb{E}_t^S[x_{t+1}] + \frac{1-\lambda^{n-2}}{1-\lambda}\mathbb{E}_t^S[x_{t+2}]$$

Hence

$$\begin{aligned}\Theta_t^{B,IRM}(n) &= -\mathbb{E}_t\left[x_{t+1} + \frac{1-\lambda^{n-2}}{1-\lambda}\mathbb{E}_{t+1}^S[x_{t+2}] - \mathbb{E}_t^S[x_{t+1}] - \frac{1-\lambda^{n-2}}{1-\lambda}\mathbb{E}_t^S[x_{t+2}]\right] \\ &= -\left[1 - \frac{\lambda k(1-\lambda^{n-2})}{1-\lambda}\right]\left[\mathbb{E}_t x_{t+1} - \mathbb{E}_t^S x_{t+1}\right]\end{aligned}$$

and therefore

$$\Theta_t^{FX,IRM}(n) = \Theta_t^{B,IRM}(n) + \Theta_t^{FX,IRM} = \left[\mathbb{E}_t x_{t+1} - \mathbb{E}_t^S x_{t+1}\right]\frac{k\lambda^{n-1}}{1-\lambda}$$

We then solve for the expressions for the interest rate misperception components of the predictability coefficients. Moreover,

$$\beta^{FX,IRM} = \frac{\mathbb{C}ov(\Theta_t^{FX,IRM}, x_t)}{\mathbb{V}ar(x_t)} \tag{34}$$

On the other hand

$$\begin{gathered}cov(x_t, \Theta_t^{FX,IRM}) = \left[1 + \frac{\lambda k}{1-\lambda}\right]cov(x_t, \mathbb{E}_t[x_{t+1}] - \mathbb{E}_t^S[x_{t+1}]) = \\ Var(x_t)\lambda\left[1 + \frac{\lambda k}{1-\lambda}\right] - \left[1 + \frac{\lambda k}{1-\lambda}\right]cov(x_t, \mathbb{E}_t^S[x_{t+1}]).\end{gathered}$$

Also

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

$$\begin{aligned}cov(x_t, \mathbb{E}_t^S[x_{t+1}]) &= Var(x_t)[\lambda k + k(1-k)\lambda^3 + k(1-k)^2\lambda^5 + \ldots] \\ &= \frac{\lambda k}{1-(1-k)\lambda^2}Var(x_t).\end{aligned}$$

Hence

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

and

$$\beta^{FX,IRM} = \left[1 + \frac{\lambda k}{1-\lambda}\right]\frac{(1-k)(\lambda - \lambda^3)}{1-(1-k)\lambda^2}.$$

The expression for $\beta^{FX,IRM}(n)$ follows similarly.

6.2 Proof of Proposition 2

We have

$$\underbrace{\Theta_t^{FX}}_{\text{Currency premium}} = \underbrace{\zeta_t^{FX}}_{\text{Risk premium differential}} + \underbrace{\mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} x_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{\infty} x_{t+1+j}\right]}_{\text{Interest rate misperception effect}}$$
$$-\underbrace{\mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} \zeta_{t+1+j}^{FX} - \mathbb{E}_t^S \sum_{j=0}^{\infty} \zeta_{t+1+j}^{FX}\right]}_{\text{Risk premium misperception effect}} + \underbrace{\mathbb{E}_t[\lim_{j\to\infty} \mathbb{E}_{t+1}^S[s_{t+j}] - \lim_{j\to\infty} \mathbb{E}_t^S[s_{t+j}]]}_{\text{Permanent component misperception effect}}$$

Condition CRP implies that the risk premium misperception effect is zero. Condition NLRM implies that the permanent component misperception effect is zero. Moreover, because each component in the expression is demeaned, condition CRP implies that risk premium differential is zero. Hence we then have

$$\underbrace{\Theta_t^{FX}}_{\text{Currency premium}} = \underbrace{\mathbb{E}_t\left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} x_{t+1+j} - \mathbb{E}_t^S \sum_{j=0}^{\infty} x_{t+1+j}\right]}_{\text{Interest rate misperception effect}}$$

Similarly, condition CRP and NLRM also imply that the interest rate misperception component drives all variation in bond premia. Now assuming condition the expressions for currency and bond premia follow directly from the same manipulations as in the proof of Proposition 2.

6.3 Proof of Proposition 3

Given conditions SE, CRP and NLRM, the currency risk premium is given by

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

Here

$$\mathbb{E}_t[x_{t+1}] = \lambda x_t$$

and

$$\mathbb{E}_t^S[x_{t+1}] = k\lambda x_t + k(1-k)\lambda^2 x_{t-1} + k(1-k)^2\lambda^3 x_{t-2} + \dots$$

Therefore

$$\Theta_t = \left[1 + \frac{\lambda k}{1-\lambda}\right]\left[\lambda(1-k)x_t - k(1-k)\lambda^2 x_{t-1} - k(1-k)^2\lambda^3 x_{t-2} - \dots\right] =$$
$$\left[1 + \frac{\lambda k}{1-\lambda}\right]\lambda(1-k)\left[x_t - k\lambda\bar{x}_{t-1}\right]$$

Hence

$$r_{t+1}^{FX} = \left[1 + \frac{\lambda k}{1-\lambda}\right]\lambda(1-k)\left[x_t - k\lambda\bar{x}_{t-1}\right] + \epsilon_{t+1}$$

This implies

$$\beta_1^{FX} = \left[1 + \frac{\lambda k}{1-\lambda}\right]\lambda(1-k).$$

and

$$\beta_2^{FX} = -\left[1 + \frac{\lambda k}{1-\lambda}\right]\lambda(1-k)k\lambda.$$

The signs are as predicted by the proposition. The proof for the bond regression is similar.

6.4 Proof of Proposition 4

Consider the regression

$$s_{t+j} - \mathbb{E}_t^S[s_{t+j}] = \alpha + \beta x_t + e_{t+j}.$$

Using the decomposition derived earlier

$$s_{t+j} - \mathbb{E}_t^S[s_{t+j}] = \mathbb{E}_{t+j}^S \sum_{h=0}^{\infty} x_{t+j+h} - \mathbb{E}_{t+j}^S \sum_{h=0}^{\infty} \zeta_{t+j+h}^{FX} + \lim_{h\to\infty} \mathbb{E}_{t+j}^S[s_{t+j+h}]$$
$$-\mathbb{E}_t^S \sum_{h=0}^{\infty} x_{t+j+h} + \mathbb{E}_t^S \sum_{h=0}^{\infty} \zeta_{t+j+h}^{FX} - \lim_{h\to\infty} \mathbb{E}_t^S[s_{t+j+h}]$$

Taking expectations under the objective measure and imposing conditions NRPM and NLRM we obtain

$$\mathbb{E}_t[s_{t+j}] - \mathbb{E}_t^S[s_{t+j}] = \mathbb{E}_t \mathbb{E}_{t+j}^S \sum_{h=0}^{\infty} x_{t+j+h} - \mathbb{E}_t^S \sum_{h=0}^{\infty} x_{t+j+h}$$

We need to evaluate

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

for $h > 0$. Recall that

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

Hence after some algebra

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

On the other hand (see the proof of proposition 1)

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

Therefore

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

The term

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

governs the sign of the coefficient. Now

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

When $h = 0$ we need to evaluate:

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

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

$$= \lambda^j \left(1 - \frac{k}{1-(1-k)\lambda^2} \right) Var(x_t) = \lambda^j \frac{(1-k)(1-\lambda^2)}{1-(1-k)\lambda^2} Var(x_t) > 0$$

Hence

$$\beta = \frac{\mathbb{C}ov(\mathbb{E}_t \mathbb{E}_{t+j}^S \sum_{h=0}^{\infty} x_{t+j+k} - \mathbb{E}_t^S \sum_{h=0}^{\infty} x_{t+j+k}, x_t)}{\mathbb{V}ar(x_t)}$$

$$= \underbrace{\lambda^j \frac{(1-k)(1-\lambda^2)}{1-(1-k)\lambda^2}}_{>0} + \sum_{h=1}^{\infty} \lambda^{h-1} \underbrace{\left[\frac{k(1-k)^j - k}{1-(1-k)\lambda^2} + 1 - (1-k)^{j+1} \right]}_{>0} > 0$$

$\square$

6.5 Formulas Left Out in Proposition 5

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

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

The conditional volatilities of the persistent components are

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

for the first two states and the common state and

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

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

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

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

Here

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

and

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

6.6 Proof of Proposition 6

The standard bond pricing equation is

$$P_t(n) = \mathbb{E}_t^S[M_{t+1}P_{t+1}(n-1)],$$

which can be expressed using our previous log notation as

$$q_t(n) = log\left(\mathbb{E}_t[exp(m_{t,t+1} + q_{t+1}(n-1))]\right),$$

Let us conjecture that this has a solution of the form

$$q_t(n) = A(n) + B_1(n)z_t + B_2(n)\bar{z}_t + A_3(n)E_t^S[z_{t+1}] + A_4(n)E_t^S[\bar{z}_{t+1}] = A(n) + \mathbf{B}'(n)\mathbf{y}_t$$

Given this conjectured form $q_{t+1}(n-1)$ and $m_{t,t+1}$ are conditionally jointly normal. Hence we obtain

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

The initial values must be such that $A(0) = B_1(0) = B_2(0) = B_3(0) = B_4(0) = 0$. Note that

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

$$\mathbb{V}ar_t^S[m_{t,t+1} + q_{t+1}(n-1)] = \sigma_\epsilon^2\varphi_t^2 + \bar{\sigma}_\epsilon^2\bar{\varphi}_t^2 + B_1(n-1)\mathbb{V}ar_t^S(z_{t+1}) + B_2(n-1)\mathbb{V}ar_t^S(\bar{z}_{t+1}) +$$
$$B_3(n-1)\mathbb{V}ar_t^S(\mathbb{E}_{t+1}^S[z_{t+2}]) + B_4(n-1)\mathbb{V}ar_t^S(\mathbb{E}_{t+1}^S[\bar{z}_{t+2}]) + 2\varphi_t B_3(n-1)Cov_t^S(\epsilon_{t+1}, \mathbb{E}_{t+1}^S[z_{t+2}]) +$$
$$2\bar{\varphi}_t B_4(n-1)Cov_t^S(\bar{\epsilon}_{t+1}, \mathbb{E}_{t+1}^S[\bar{z}_{t+2}])$$

So we obtain the following equation

$$A(n) + \mathbf{B}'(n)\mathbf{y}_t = -logR - z_t - \bar{z}_t + A(n-1) + \mathbf{B}'(n-1)\mathbb{E}_t^S[\mathbf{y}_{t+1}] + \frac{1}{2}B_1(n-1)^2\mathbb{V}ar_t^S(z_{t+1}) +$$
$$\frac{1}{2}B_2(n-1)^2\mathbb{V}ar_t^S(\bar{z}_{t+1}) + \frac{1}{2}B_3(n-1)^2\mathbb{V}ar_t^S(\mathbb{E}_{t+1}^S[z_{t+2}]) + \frac{1}{2}B_4(n-1)^2\mathbb{V}ar_t^S(\mathbb{E}_{t+1}^S[\bar{z}_{t+2}]) +$$
$$\varphi_t B_3(n-1)Cov_t^S(\epsilon_{t+1}, \mathbb{E}_{t+1}^S[z_{t+2}]) + \bar{\varphi}_t B_4(n-1)Cov_t^S(\bar{\epsilon}_{t+1}, \mathbb{E}_{t+1}^S[\bar{z}_{t+2}]) +$$
$$B_1(n-1)B_3(n-1)Cov_t^S(z_{t+1}, \mathbb{E}_{t+1}^S[z_{t+2}]) + B_2(n-1)B_4(n-1)Cov_t^S(\bar{z}_{t+1}, \mathbb{E}_{t+1}^S[\bar{z}_{t+2}])$$

Recall that

$$\varphi_t = \varphi_0 + \varphi_1 z_t + \varphi_2\bar{z}_t \quad \bar{\varphi}_t = \bar{\varphi}_0 + \bar{\varphi}_1 z_t + \bar{\varphi}_2\bar{z}_t$$

Hence we have

$$A(n) = A(n-1) - logR + \frac{1}{2}B_1(n-1)^2\mathbb{V}ar_t^S(z_{t+1}) +$$
$$\frac{1}{2}B_2(n-1)^2\mathbb{V}ar_t^S(\bar{z}_{t+1}) + \frac{1}{2}B_3(n-1)^2\mathbb{V}ar_t^S(\mathbb{E}_{t+1}^S[z_{t+2}]) + \frac{1}{2}B_4^2(n-1)\mathbb{V}ar_t^S(\mathbb{E}_{t+1}^S[\bar{z}_{t+2}]) +$$
$$\varphi_0 B_3(n-1)Cov_t^S(\epsilon_{t+1}, \mathbb{E}_{t+1}^S[z_{t+2}]) + \bar{\varphi}_0 B_4(n-1)Cov_t^S(\bar{\epsilon}_{t+1}, \mathbb{E}_{t+1}^S[\bar{z}_{t+2}]) +$$
$$B_1(n-1)B_3(n-1)Cov_t^S(z_{t+1}, \mathbb{E}_{t+1}^S[z_{t+2}]) + B_2(n-1)B_4(n-1)Cov_t^S(\bar{z}_{t+1}, \mathbb{E}_{t+1}^S[\bar{z}_{t+2}])$$

$$B_1(n) = -1 + \varphi_1 B_3(n-1) Cov_t^S(\epsilon_{t+1}, \mathbb{E}_{t+1}^S[z_{t+2}]) + \bar{\varphi}_1 B_4(n-1) Cov_t^S(\bar{\epsilon}_{t+1}, \mathbb{E}_{t+1}^S[\bar{z}_{t+2}])$$

$$B_2(n) = -1 + \varphi_2 B_3(n-1) Cov_t^S(\epsilon_{t+1}, \mathbb{E}_{t+1}^S[z_{t+2}]) + \bar{\varphi}_2 B_4(n-1) Cov_t^S(\bar{\epsilon}_{t+1}, \mathbb{E}_{t+1}^S[\bar{z}_{t+2}])$$

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

The variance and covariance terms are constant. They are given by

$$Cov_t^S(\epsilon_{t+1}, \mathbb{E}_{t+1}^S[z_{t+2}]) = Cov_t^S(\epsilon_{t+1}, \lambda k z_{t+1} + (1-k)\lambda \mathbb{E}_t^S[z_{t+1}]) = \lambda k \sigma_\epsilon^2$$

and similarly

$$Cov_t^S(\bar{\epsilon}_{t+1}, \mathbb{E}_{t+1}^S[\bar{z}_{t+2}]) = \bar{\lambda}\bar{k}\bar{\sigma}_\epsilon^2$$

.

$$\mathbb{V}ar_t^S(z_{t+1}) = \sigma^2 + \sigma_v^2 \quad \mathbb{V}ar_t^S(\bar{z}_{t+1}) = \bar{\sigma}^2 + \bar{\sigma}_v^2$$

$$\mathbb{V}ar_t^S(\mathbb{E}_{t+1}^S[z_{t+2}]) = \lambda k(\sigma^2 + \sigma_v^2) \quad \mathbb{V}ar_t^S(\mathbb{E}_{t+1}^S[\bar{z}_{t+2}]) = \bar{\lambda}\bar{k}(\bar{\sigma}^2 + \bar{\sigma}_v^2)$$

Note that symmetry implies

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

The Rational Case The rational model is a special case of the above model. Here the solution is

$$q_t(n) = A_r(n) + B_{1,r}(n) z_t + B_{2,r}(n) \bar{z}_t$$

The coefficients can also be solved from

$$A_r(n) + B_{1,r}(n) z_t + B_{2,r} \bar{z}_t = -logR - z_t - \bar{z}_t + A_r(n-1) + B_{1,r}(n-1)\lambda z_t + B_{2,r}(n-1)\bar{\lambda}\bar{z}_t +$$
$$+ \frac{1}{2} B_{1,r}(n-1)^2 \sigma_\epsilon^2 + \frac{1}{2} B_{2,r}(n-1)^2 \bar{\sigma}_\epsilon^2 + \varphi_t B_{1,r}(n-1)\sigma_\epsilon^2 + \bar{\varphi}_t B_{2,r}(n-1)\bar{\sigma}_\epsilon^2$$

So that we have the solution

$$A_r(n) =$$
$$A_r(n-1) - logR + \frac{1}{2}B_{1,r}(n-1)^2\sigma_\epsilon^2 + \frac{1}{2}B_{2,r}(n-1)^2\bar{\sigma}_\epsilon^2 + \varphi_0 B_{1,r}(n-1)\sigma_\epsilon^2 + \bar{\varphi}_0 B_{2,r}(n-1)\bar{\sigma}_\epsilon^2$$

$$B_{1,r}(n) = -1 + B_{1,r}(n-1)\lambda + \varphi_1 B_{1,r}(n-1)\sigma_\epsilon^2 + \bar{\varphi}_1 B_{2,r}(n-1)\bar{\sigma}_\epsilon^2$$

$$B_{2,r}(n) = -1 + B_{2,r}(n-1)\bar{\lambda} + \varphi_2 B_{1,r}(n-1)\sigma_\epsilon^2 + \bar{\varphi}_2 B_{2,r}(n-1)\bar{\sigma}_\epsilon^2$$

Note that here

$$q_t(n)^* - q_t(n) = B_{1,r}(n)x_t$$

6.7 Closed Form Solutions for the Predictability Coefficients in the Affine Model

We now derive analytical expressions for all the predictability coefficients in the context of our affine model. These closed form expressions greatly simplify and speed up model estimation. The relative spread is

$$-\frac{B_1(n)}{n}x_t - \frac{B_3(n)}{n}\mathbb{E}_t^S[x_{t+1}] - x_t = -\left(\frac{B_1(n)}{n} + 1\right)x_t - \frac{B_3(n)}{n}\mathbb{E}_t^S[x_{t+1}]$$

And expected bond excess return is

$$\mathbb{E}_t[q_{t+1}(n-1)^* - q_{t+1}(n-1)] - (q_t(n)^* - q_t(n)) - x_t =$$
$$(B_1(n-1)\lambda - B_1(n) - 1)x_t + (B_3(n-1)\lambda - B_3(n))\mathbb{E}_t^S[x_{t+1}]$$

Note

$$Cov(\mathbb{E}_t[q_{t+1}(n-1)^* - q_{t+1}(n-1)] - (q_t(n)^* - q_t(n)) - x_t, x_t) =$$
$$(B_1(n-1)\lambda - B_1(n) - 1)\mathbb{V}ar(x_{t+1}) + (B_3(n-1)\lambda - B_3(n))Cov(x_{t+1}, \mathbb{E}_t^S[x_{t+1}])$$

When regressing bond excess returns on relative short rates we then obtain a predictability coefficient of

$$\frac{Cov(\mathbb{E}_t[q_{t+1}(n-1)^* - q_{t+1}(n-1)] - (q_t(n)^* - q_t(n)) - x_t, x_t)}{\mathbb{V}ar(x_t)} =$$
$$(B_1(n-1)\lambda - B_1(n) - 1) + (B_3(n-1)\lambda - B_3(n))\frac{\lambda k}{1-(1-k)\lambda^2}$$

Moreover, for the spread we have

$$\begin{gathered}
Cov(\mathbb{E}_t[q_{t+1}(n-1)^* - q_{t+1}(n-1)] - (q_t(n)^* - q_t(n)) - x_t, \\
-\left(\frac{B_1(n)}{n}+1\right)x_t - \frac{B_3(n)}{n}\mathbb{E}_t^S[x_{t+1}]) = \\
-\left(\frac{B_1(n)}{n}+1\right)(B_1(n-1)\lambda - B_1(n) - 1)\mathbb{V}ar(x_{t+1}) \\
-\left[\left(\frac{B_1(n)}{n}+1\right)(B_3(n-1)\lambda - B_3(n)) + \frac{B_3(n)}{n}(B_1(n-1)\lambda - B_1(n) - 1)\right]Cov(x_{t+1}, \mathbb{E}_t^S[x_{t+1}]) \\
-\frac{B_3(n)}{n}(B_3(n-1)\lambda - B_3(n))\mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}])
\end{gathered}$$

Here

$$\mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}]) = \frac{k^2\lambda^2 + 2k(1-k)\lambda^2\frac{\lambda k}{1-(1-k)\lambda^2}}{1-(1-k)^2\lambda^2}\mathbb{V}ar(x_t)$$

the variance of the spread is

$$\begin{gathered}
\mathbb{V}ar(-\left(\frac{B_1(n)}{n}+1\right)x_t - \frac{B_3(n)}{n}\mathbb{E}_t^S[x_{t+1}]) = \\
\left(\frac{B_1(n)}{n}+1\right)^2\mathbb{V}ar(x_t) + 2\left(\frac{B_1(n)}{n}+1\right)\frac{B_3(n)}{n}Cov(\mathbb{E}_t^S[x_{t+1}], x_t) + \frac{B_3(n)^2}{n^2}\mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}])
\end{gathered}$$

and the predictability coefficient is ratio of the above covariance terms. Similarly the variance of n maturity relative yield is

$$\begin{gathered}
\mathbb{V}ar(-\frac{B_1(n)x_t}{n} - \frac{B_3(n)\mathbb{E}_t^S[x_{t+1}]}{n}) = \\
\frac{B_1(n)^2}{n^2}\mathbb{V}ar(x_t) + 2\frac{B_1(n)B_3(n)}{n^2}Cov(\mathbb{E}_t^S[x_{t+1}], x_t) + \frac{B_3(n)^2}{n^2}\mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}])
\end{gathered}$$

Next consider currencies. Using our previous notation we have

$$\zeta_t^{FX} = -\frac{\bar{\sigma}_\epsilon^2\bar{\varphi}_t^2}{2} - \frac{\sigma_\epsilon^2\varphi_t^2}{2} + \frac{\bar{\sigma}_\epsilon^2\bar{\varphi}_t^{*2}}{2} + \frac{\sigma_\epsilon^2\varphi_t^{*2}}{2} =$$
$$\bar{\sigma}_\epsilon^2(\bar{\varphi}_0\bar{\varphi}_1(z_t - z_t^*) + \bar{\varphi}_1^2(z_t^2 - z_t^{*2}) + 2\bar{\varphi}_1\bar{\varphi}_2\bar{z}_t(z_t - z_t^*))+$$
$$\sigma_\epsilon^2(\varphi_0\varphi_1(z_t - z_t^*) + \varphi_1^2(z_t^2 - z_t^{*2}) + 2\varphi_1\varphi_2\bar{z}_t(z_t - z_t^*))$$

Moreover

$$Cov(\zeta_t^{FX}, x_t) = (-\bar{\sigma}_\epsilon^2\bar{\varphi}_0\bar{\varphi}_1 - \sigma_\epsilon^2\varphi_0\varphi_1)\mathbb{V}ar(x_t)$$

and

$$Cov(\mathbb{E}_t^S\zeta_{t+j}^{FX}, x_t) = (-\bar{\sigma}_\epsilon^2\bar{\varphi}_0\bar{\varphi}_1 - \sigma_\epsilon^2\varphi_0\varphi_1)\mathbb{C}ov(\mathbb{E}_t^S[x_{t+j}], x_t) = \lambda^{j-1}\frac{\lambda k}{1-(1-k)\lambda^2}Var(x_t)$$

and

$$Cov(\zeta_{t+1}^{FX}, x_t) = \lambda(-\bar{\sigma}_\epsilon^2\bar{\varphi}_0\bar{\varphi}_1 - \sigma_\epsilon^2\varphi_0\varphi_1)$$

$$Cov(\mathbb{E}_{t+1}^S\zeta_{t+j+1}^{FX}, x_t) = \lambda^{j-1}(-\bar{\sigma}_\epsilon^2\bar{\varphi}_0\bar{\varphi}_1 - \sigma_\epsilon^2\varphi_0\varphi_1)Cov(\mathbb{E}_{t+1}^S x_{t+2}, x_t) =$$
$$\lambda^{j-1}(-\bar{\sigma}_\epsilon^2\bar{\varphi}_0\bar{\varphi}_1 - \sigma_\epsilon^2\varphi_0\varphi_1)(\lambda k + (1-k)\lambda\frac{\lambda k}{1-(1-k)\lambda^2})$$

$$s_t = \sum_{j=0}^{\infty}\mathbb{E}_t^S[m_{t+j,t+j+1} - m_{t+j,t+j+1}^*] + \lim_{j\to\infty}\mathbb{E}_t^S[s_{t+j}]$$

Hence we obtain a predictability coefficient related to x_t of

$$\beta_{FX,x} = \beta^{IRM} +$$
$$(-\bar{\sigma}_\epsilon^2\bar{\varphi}_0\bar{\varphi}_1 - \sigma_\epsilon^2\varphi_0\varphi_1)\left(1 - \lambda - \frac{1}{1-\lambda}(\lambda k + (1-k)\lambda\frac{\lambda k}{1-(1-k)\lambda^2}) + \frac{1}{1-\lambda}\frac{\lambda k}{1-(1-k)\lambda^2}\right)$$

We now need to solve for the predictability coefficient related to spread. Here note

$$Cov(\zeta_t^{FX}, \mathbb{E}_t^S[x_{t+1}]) = (-\bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 - \sigma_\epsilon^2 \varphi_0 \varphi_1) \frac{\lambda k}{1-(1-k)\lambda^2}$$

$$Cov(\mathbb{E}_t^S \zeta_{t+j}^{FX}, \mathbb{E}_t^S[x_{t+1}]) = (-\bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 - \sigma_\epsilon^2 \varphi_0 \varphi_1) \lambda^{j-1} \mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}]) = \\ \lambda^{j-1}(-\bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 - \sigma_\epsilon^2 \varphi_0 \varphi_1) \mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}])$$

$$Cov(\zeta_{t+1}^{FX}, \mathbb{E}_t^S[x_{t+1}]) = \lambda(-\bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 - \sigma_\epsilon^2 \varphi_0 \varphi_1) \frac{\lambda k}{1-(1-k)\lambda^2} \mathbb{V}ar(x_t)$$

$$Cov(\mathbb{E}_{t+1}^S \zeta_{t+j+1}^{FX}, \mathbb{E}_t^S[x_{t+1}]) = \lambda^{j-1} Cov(\mathbb{E}_{t+1}^S \zeta_{t+2}^{FX}, \mathbb{E}_t^S[x_{t+1}]) = \\ \lambda^{j-1}(-\bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 - \sigma_\epsilon^2 \varphi_0 \varphi_1) Cov(\mathbb{E}_{t+1}^S[x_{t+2}], \mathbb{E}_t^S[x_{t+1}]) = \\ \lambda^{j-1}(-\bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 - \sigma_\epsilon^2 \varphi_0 \varphi_1) \left(\lambda^2 k \frac{\lambda k}{1-(1-k)\lambda^2} \mathbb{V}ar(x_t) + (1-k)\lambda \mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}]) \right)$$

$$Cov(\Theta_t^{FX,RPM}, \mathbb{E}_t^S[x_{t+1}]) = Cov(-\mathbb{E}_t \left[\mathbb{E}_{t+1}^S \sum_{j=0}^{\infty} \zeta_{t+1+j}^{FX} - \mathbb{E}_t^S \sum_{j=0}^{\infty} \zeta_{t+1+j}^{FX} \right], \mathbb{E}_t^S[x_{t+1}]) = \\ (\bar{\sigma}_\epsilon^2 \bar{\varphi}_0 \bar{\varphi}_1 + \sigma_\epsilon^2 \varphi_0 \varphi_1)(a_1 + a_2 - a_3)$$

Here

$$a_1 = \frac{\lambda^2 k}{1-(1-k)\lambda^2} \mathbb{V}ar(x_t)$$

$$a_2 = \frac{1}{1-\lambda} \left(\lambda^2 k \frac{\lambda k}{1-(1-k)\lambda^2} \mathbb{V}ar(x_t) + (1-k)\lambda \mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}]) \right)$$

$$a_3 = \frac{1}{1-\lambda} \mathbb{V}ar(\mathbb{E}_t^S[x_{t+1}])$$

We also have

$$Cov(\mathbb{E}_t\left[\mathbb{E}^S_{t+1}\sum_{j=0}^{\infty}x_{t+1+j}-\mathbb{E}^S_t\sum_{j=0}^{\infty}x_{t+1+j},\mathbb{E}^S_t[x_{t+1}]\right]=$$
$$\left[1+\frac{\lambda k}{1-\lambda}\right]\left[\mathbb{E}_t x_{t+1}-\mathbb{E}^S_t x_{t+1}\right]Cov(\mathbb{E}_t x_{t+1}-\mathbb{E}^S_t x_{t+1},\mathbb{E}^S_t x_{t+1})=$$
$$\left[1+\frac{\lambda k}{1-\lambda}\right](\lambda Cov(x_t,\mathbb{E}^S_t[x_{t+1}])-\mathbb{V}ar(\mathbb{E}^S_t[x_{t+1}]))$$

Hence for the spread we obtain

$$Cov(\mathbb{E}^S_t[s_{t+1}]-s_t+x_t,-\left(\frac{B_1(n)}{n}+1\right)x_t-\frac{B_3(n)}{n}\mathbb{E}^S_t[x_{t+1}])=$$
$$-\left(\frac{B_1(n)}{n}+1\right)\beta_{FX,x}\mathbb{V}ar(x_t)+$$
$$-\frac{B_3(n)}{n}\left(\left[1+\frac{\lambda k}{1-\lambda}\right](\lambda Cov(x_t,\mathbb{E}^S_t[x_{t+1}])-\mathbb{V}ar(\mathbb{E}^S_t[x_{t+1}]))\right)+$$
$$-\frac{B_3(n)}{n}\left(Cov(\zeta^{FX}_{t+1},\mathbb{E}^S_t[x_{t+1}])+Cov(\Theta^{FX,RPM}_t,\mathbb{E}^S_t[x_{t+1}])\right)$$

Again the variance of the spread is

$$\mathbb{V}ar(-\left(\frac{B_1(n)}{n}+1\right)x_t-\frac{B_3(n)}{n}\mathbb{E}^S_t[x_{t+1}])=$$
$$\left(\frac{B_1(n)}{n}+1\right)^2\mathbb{V}ar(x_t)+2\left(\frac{B_1(n)}{n}+1\right)\frac{B_3(n)}{n}Cov(\mathbb{E}^S_t[x_{t+1}],x_t)+\frac{B_3(n)^2}{n^2}\mathbb{V}ar(\mathbb{E}^S_t[x_{t+1}])$$

Predictability Coefficients for the Rational Model

The predictability coefficients for the rational model are obtained as a special case of the above coefficients. Here

$$q_t(n)^*-q_t(n)=B_{1,r}(n)x_t$$

The spread is

$$q_t(n)^*-q_t(n)=-B_{1,r}(n)x_t/n-x_t=-(B_{1,r}(n)/n+1)x_t$$

And expected bond excess return is

$$\mathbb{E}_t[q_{t+1}(n-1)^* - q_{t+1}(n-1)] - (q_t(n)^* - q_t(n)) - x_t = (B_{1,r}(n-1)\lambda - B_{1,r}(n) - 1)x_t$$

Hence the predictability coefficient is

$$\frac{Cov(\mathbb{E}_t[q_{t+1}(n-1)^* - q_{t+1}(n-1)] - (q_t(n)^* - q_t(n)) - x_t, x_t)}{Var(x_t)} = \\ (B_{1,r}(n-1)\lambda - B_{1,r}(n) - 1)$$

And for the spread:

$$\frac{Cov(\mathbb{E}_t[q_{t+1}(n-1)^* - q_{t+1}(n-1)] - (q_t(n)^* - q_t(n)) - x_t, -(B_{1,r}(n)/n+1)x_t)}{Var(-(B_{1,r}(n)/n+1)x_t)} = \\ -\frac{B_{1,r}(n-1)\lambda - B_{1,r}(n) - 1}{B_{1,r}(n)/n+1}$$

In the case of rational expectations we have

$$\mathbb{E}_t[s_{t+1}] - s_t + x_t = \frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^2}{2} + \frac{\sigma_\epsilon^2 \varphi_t^2}{2} - \frac{\bar{\sigma}_\epsilon^2 \bar{\varphi}_t^{*2}}{2} - \frac{\sigma_\epsilon^2 \varphi_t^{*2}}{2} = \\ \bar{\sigma}_\epsilon^2(\bar{\varphi}_0\bar{\varphi}_1(z_t - z_t^*) + \bar{\varphi}_1^2(z_t^2 - z_t^{*2}) + 2\bar{\varphi}_1\bar{\varphi}_2\bar{z}_t(z_t - z_t^*)) + \\ \sigma_\epsilon^2(\varphi_0\varphi_1(z_t - z_t^*) + \varphi_1^2(z_t^2 - z_t^{*2}) + 2\varphi_1\varphi_2\bar{z}_t(z_t - z_t^*))$$

The the predictability coefficients for currencies are:

$$\frac{Cov(\mathbb{E}_t[s_{t+1}] - s_t + x_t, x_t)}{Var(x_t)} = -\bar{\sigma}_\epsilon^2\bar{\varphi}_0\bar{\varphi}_1 - \sigma_\epsilon^2\varphi_0\varphi_1$$

and for the spread:

$$\frac{Cov(\mathbb{E}_t[s_{t+1}] - s_t + x_t, -(B_{1,r}(n)/n+1)x_t)}{Var(-(B_{1,r}(n)/n+1)x_t)} = \\ \frac{\bar{\sigma}_\epsilon^2\bar{\varphi}_0\bar{\varphi}_1 + \sigma_\epsilon^2\varphi_0\varphi_1}{B_{1,r}(n)/n+1}$$

Moreover, the variance of n maturity relative yield is $\frac{B_{1,r}(n)^2}{n^2}$.

6.8 On Estimating k

This section derives the slope coefficient in the regression where forecast errors are explained by forecast revisions. Similarly to Coibion and Gorodnichenko (2012) we have

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

$$\mathbb{E}_t[x_{t+1}] = \lambda x_t.$$

Multiplying the first expression by λ^{j-1}:

$$\mathbb{E}_t^S[x_{t+j}] = (1-k)\mathbb{E}_{t-1}^s[x_{t+j}] + k\mathbb{E}[x_{t+j}],$$

where we used the property $\mathbb{E}_t^S[x_{t+j}] = \lambda^{j-1}\mathbb{E}_t^S[x_{t+1}]$. From this it follows that

$$k(\mathbb{E}[x_{t+j}] - \mathbb{E}^S[x_{t+j}]) = (1-k)(\mathbb{E}_t^S[x_{t+j}] - \mathbb{E}_{t-1}^S[x_{t+j}]).$$

Hence

$$x_{t+j} - \mathbb{E}^S[x_{t+j}] = \frac{1-k}{k}(\mathbb{E}_t^S[x_{t+j}] - \mathbb{E}_{t-1}^S[x_{t+j}]) + u_{t+j},$$

where u_{t+j} is zero mean and orthogonal to time t information. Hence $\beta_1^{FR} = \frac{1-k}{k}$ and $\beta_0^{FR} = 0$.

6.9 Robustness Checks for Empirical Analysis

We conduct several robustness checks for our results. First, some authors such as Engel (2016) voluntarily leave the period after the financial crisis out from the sample due to possible changes in the driving forces of currencies. Similarly this period might be extraordinary for the bond market due to low interest rates and unconventional monetary policies. Tables 9, 10 and 11 replicate tables 2, 3 and 4 but now excluding the period after 2008. Excluding this period does not alter the key results: rather many of results become stronger. The results in the after 2008 subsample are somewhat weaker and mostly not statistically significant. However, the sample period is fairly short. As mentioned before, many of our results also become stronger if we omit Japan, where interest rates have been very low during most of the sample period.

	$\hat{\lambda}_0$	$s.e$	$\hat{\lambda}_1$	$s.e$	R^2
panel			0.982***	0.009	0.978
CAN	0.024	0.034	0.956***	0.019	0.934
GER	-0.005	0.032	0.992***	0.015	0.991
JAP	-0.078	0.029	0.979***	0.008	0.990
NOR	0.049	0.072	0.986***	0.021	0.974
SWE	-0.008	0.043	0.986***	0.016	0.981
CH	-0.004	0.046	0.986***	0.016	0.962
UK	0.045	0.046	0.972***	0.021	0.970

Table 9 shows the results from regressing monthly 3 month yield differential (foreign minus US) on its first lag excluding the sample period after 2008. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

	$\hat{\beta}_0$	$s.e$	$\hat{\beta}_1$	$s.e$	R^2	*implied k*
panel			1.648**	0.500	0.061	0.38
CAN	0.333	0.263	0.049	0.362	0.001	0.95
GER	0.538**	0.258	2.826***	0.681	0.152	0.26
JAP	0.432	0.311	1.936***	0.656	0.081	0.34
NOR	0.664	0.511	3.315***	0.979	0.172	0.23
SWE	-0.150	0.396	2.103***	0.800	0.092	0.32
CH	0.593*	0.339	1.766**	0.724	0.073	0.36
UK	0.412	0.266	0.580	0.480	0.010	0.63

Table 10 shows the results from regressing the difference in forecast error (foreign minus US) from forecasting spot 3 month 12 months ahead on the difference in forecast revisions excluding the sample period after 2008. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

PANEL A: Bond Returns Before 2008										
	3 month rate					yield slope				
	$\hat{\beta}_0$	s.e	$\hat{\beta}_1(3)$	s.e	R^2	$\hat{\beta}_0$	s.e	$\hat{\beta}_1$	s.e	R^2
panel			-1.322***	0.310	0.022			1.421**	0.530	0.012
AUS	0.002**	0.002	-1.382* *	0.600	0.030	0.001	0.001	2.333**	1.074	0.026
CAN	0.002*	0.001	-1.989***	0.562	0.051	0.001*	0.001	2.805***	0.777	0.050
GER	-0.002*	0.001	-0.806**	0.400	0.007	-0.001	0.001	0.764	0.529	0.005
JAP	-0.003**	0.002	-1.367***	0.512	0.089	-0.000	0.001	0.621	0.508	0.002
NZL	0.002	0.002	-1.675***	0.650	0.013	-0.001	0.001	0.975	0.873	0.004
SWE	0.002	0.001	-1.195**	0.616	0.025	0.001	0.001	1.579	1.070	0.022
CH	-0.003***	0.001	-1.163**	0.553	0.011	-0.001	0.001	1.198*	0.655	0.011
UK	0.002	0.002	-1.544***	0.667	0.017	-0.001	0.001	0.892	0.815	0.003

PANEL B: Currency Returns Before 2008										
	3 month rate					yield slope				
	$\hat{\beta}_0$	s.e	$\hat{\beta}_1(3)$	s.e	R^2	$\hat{\beta}_0$	s.e	$\hat{\beta}_1$	s.e	R^2
panel			1.758**	0.515	0.026			-2.834**	0948	0.020
AUS	-0.001	0.003	1.527**	0.621	0.027	-0.000	0.002	-2.999***	1.200	0.033
CAN	0.002	0.001	1.424**	0.642	0.024	0.003**	0.001	-2.112***	0.840	0.025
GER	0.008***	0.003	-.285	1.323	0.000	0.007**	0.003	-1.088	1.748	0.003
JAP	0.001**	0.003	4.253***	1.115	0.066	-0.003*	0.002	-4.617***	1.133	0.045
NZL	-0.004	0.003	2.386***	0.370	0.100	-0.005*	0.003	-4.926***	0.602	0.125
SWE	-0.002	0.002	0.942	1.107	0.011	-0.001	0.002	-0.409	1.649	0.001
CH	0.004*	0.002	1.958*	1.206	0.016	-0.000	0.002	-2.781*	1.518	0.022
UK	0.001	0.002	1.648*	0.997	0.014	0.001	0.002	-2.507	1.596	0.016

Table 11 shows the results from regressing the relative bond (foreign minus US) and currency returns on short-rate and yield spread differences excluding the sample after 2008. The standard errors of the panel regression are calculated using the (Driscoll and Kraay, 1998) methodology with 13 lags, which corrects for heteroskedasticity, serial correlation, and cross-equation correlation. The standard errors for individual regressions are corrected for heteroskedasticity and autocorrelation (Newey and West, 1987). *, ** and *** denote significance at 10%, 5% and 1% levels respectively.

Coefficient	Data	β^{IRM} under SE	Model SE+CRP+NLRM
β, LHS: Currret, RHS: x_t	1.489	0.948 (63 %)	0.948
β, LHS: Currret, RHS: yield spread	-1.943	NA	-1.35
β, LHS: $r^{FX}_{t+1}(120)$, RHS: x_t	0.23	0.04 (17 %)	0.04
β, LHS: $r^{FX}_{t+1}(120)$, RHS: yield spread	-0.69	NA	0.05
β, LHS: Bondret, RHS: x_t	-1.259	-0.99 (79 %)	-0.99
β, LHS: Bondret, RHS: yield spread	1.250	NA	1.40
Volatility ratio, 10 year rate, 3 month rate	0.67	NA	0.30

Table 12 shows key statistics measured from the data (panel regressions) as well as those predicted by the model, AR(2) -version of the model.

Note that our assumptions imply that under correct beliefs $k = 1$. However, in theory some underweighting might be statistically optimal e.g. due to noisy observations. We now test this assumption of the model. Using the actual short rate process we obtain a panel estimate $k \approx 0.983$.[19] Moreover, $k = 1$ clearly cannot be rejected. Therefore assuming $k = 1$ under correct beliefs seems empirically reasonable.

The baseline model assumes an AR(1)-process for the short rate differential under the objective measure. We now estimate a sticky expectations AR(2)-version of the model. Table 12 replicates 5 but now under the assumption that the true short rate differential process is AR(2). These are obtained using simulations as the AR(2)-version does not allow for simple closed-form expressions. One can see that the sticky expectations AR(2) -model gives a somewhat more accurate results for the predictability coefficients. However, it understates relative long rate volatility.

[19] This can be estimated either using the above regression procedure or maximum likelihood.