International yield curves and currency puzzles*

Mikhail Chernov[†] and Drew Creal[‡]

First Draft: June 21, 2017
This Draft: May 9, 2019

Abstract

The depreciation rate is often computed as the ratio of foreign and domestic pricing kernels. Using bond prices alone to estimate these kernels leads to currency puzzles: the inability of models to match violations of uncovered interest parity and the volatility of exchange rates. That happens because of the *FX bond disconnect*, the inability of bonds to span exchange rates. This view of the puzzles is distinct from market incompleteness. Incorporating exchange rates into estimation of yield curve models helps with resolving the puzzles. That approach also allows one to relate news about the cross-country differences between international yields to news about currency risk premiums.

JEL Classification Codes: F31, G12, G15.

Keywords: exchange rates, bond valuation, FX disconnect, affine models.

---

* We thank Michael Bauer, Magnus Dalhquist, Greg Duffee, Bernard Dumas, Federico Gavazzoni, Jeremy Graveline, Valentin Haddad, Sebastian Infante, Rob Richmond, Chris Telmer, Irina Zviadadze for comments on earlier drafts and participants in the seminars and conferences sponsored by Adam Smith Asset Pricing Conference 2019, Santiago Finance Workshop 2018, SFS Cavalcade 2019, McGill University. The latest version is available at

https://sites.google.com/site/mbchernov/CC_DIY_latest.pdf.

† Anderson School of Management, UCLA, NBER, and CEPR; mikhail.chernov@anderson.ucla.edu.

‡ Department of Economics, University of Notre Dame; dcreal@nd.edu.

1 Introduction

The asset market view of exchange rates, whereby the depreciation rate is computed as the ratio of foreign and domestic pricing kernels, has become a dominant paradigm in financial economics following the influential work of Backus, Foresi, and Telmer (2001) (B/F/T hereafter). Thus, empirical applications require estimates of the pricing kernels, and looking to foreign and domestic bond prices as the relevant source of information appears to be a natural step. Indeed, that was pursued by B/F/T, among many others.

A typical finding is that the variation in depreciation rates, inferred via the asset market view, has little to do with that of the observed ones, referred to as the *FX volatility anomaly* (Brandt and Santa-Clara, 2002), and that the inferred depreciation rates cannot replicate the *FX forward premium anomaly,* that is, the well-documented violations of the uncovered interest parity (UIP) hypothesis (B/F/T).

In this paper we argue that one must use information about depreciation rates jointly with bond prices in order to understand the cross-country differences between bonds and their connection to exchange rates. That is because, as we show, depreciation rates are not spanned by bonds. As a result, one cannot use information in bonds alone to infer the dynamics of depreciation rates. This focus is distinct from exploration of market incompleteness as a potential source of FX anomalies (e.g., Lustig and Verdelhan, 2015). International business cycle models could imply lack of FX spanning with bonds in economies with both complete and incomplete financial markets, depending on the transmission of output shocks and their interaction with relative prices of imports and exports.

The flip side of using bonds to infer pricing kernels is an asset-pricing exercise in valuation of international bonds. Typically, the literature on this topic focuses on similarities between the different countries by modeling global/US and local/foreign factors and quantifying their contribution to the overall variation of the curves. The conclusion is that common variation is the major driver of interest rates. The sole focus on commonalities is surprising given the aforementioned prominent connection between bonds and exchange rates in the currency literature.

Indeed, we show that the differences in domestic and foreign bonds must be related to depreciation rates. Specifically, cross-country differences between yields reflect expected future depreciation rates and the associated currency risk premiums. Further, cross-country differences between bond risk premiums reflect currency risk premiums. We demonstrate that in our sample of three countries (USA, Germany, and UK) the second and third principal components of the joint set of three yield curves are approximated by these differences. This evidence suggests an important role for depreciation rates and the associated currency risk premiums in understanding international yield curves.

We explore practical implications of these ideas via the joint modeling of bonds and currencies using the no-arbitrage affine framework as in B/F/T. This approach allows us to estimate the pricing kernels expressed in USD using data on U.S. and foreign bond prices, and on exchange rates (referred to as the WFX approach, estimation With FX rates). This is in contrast to the predominant approach that estimates the pricing kernel denominated in currency of a given country using data on bonds of the same country that are denominated in the currency of that country (referred to as the NFX approach, No FX rates used in estimation).

We show that the two approaches match yields equally well, consistent with exchange rates that are unspanned by bonds. However, the exchange rate implied by the NFX approach is grossly misspecified. Its behavior is in line with findings reported by previous studies. In contrast, the WFX approach implies realistic exchange rate behavior. In particular, we can match all the FX moments discussed by B/F/T.

The NFX approach does not allow one to explore how currency risk premiums connect bond yields and bond risk premiums of different countries. We use the WFX model to interpret the differences between the international yield curves. We decompose news about currency risk premium into news about expected depreciation rate and cross-country bond yield differential. The latter contributes very little at short horizon and the contribution grows to about 50% at long horizons. In contrast, if one were to use the NFX model and combine it with the UIP regression to get a sensible estimate of the currency risk premium, the contribution of the yield differential would be about 80% regardless of horizon.

The main lesson from our empirical study is that augmenting the set of assets by exchange rates makes a big difference in the implications for the analysis of international asset markets. In particular, this suggests that a rich collection of international bonds does not complete the markets. Following the literature on the FX macro disconnect, we refer to this finding as the *FX bond disconnect*. Market incompleteness with respect to bonds should be a starting point for any equilibrium model of exchange rates.

Our paper is related to several research themes. We do not provide a grand literature review. Instead, in an attempt to offer clarity, we describe the related work when we cover an appropriate subtopic.

2 Preliminaries

In this section we introduce definitions and notation used throughout the paper. We specifically highlight differences between two types of objects. First, we draw a distinction between the true and estimated pricing kernels. The emphasis is needed because, depending on the data used for estimation, one might not be able to estimate the true pricing kernel even if its functional form is known. Second, we distinguish market completeness and hidden factors. Informally, the two concepts sound similar, but they are not equivalent. This lack of equivalence is important for our analysis because it allows us to avoid taking a stand on whether markets are complete.

2.1 Bonds and currencies

Suppose $M_{t,t+i}$ is an i-period pricing kernel expressed in USD. Then the USD-denominated value of any zero-coupon bond of maturity n is

$$P_t^n = E_t(M_{t,t+n} \cdot C_{t,t+n}),$$

where $C_{t,t+i}$ is the cash flow growth between time t and $t+i$. If the bond is issued in USD, then $C_{t,t+i} = 1$; we denote its price by Q_t^n and its yield is $y_t^n = -n^{-1} \log Q_t^n \equiv -n^{-1} q_t^n$. If the bond is issued in foreign currency, then $C_{t,t+i} = S_{t+i}/S_t$ with S_t representing the value of one unit of foreign currency in USD; we denote the foreign bond price by $\widehat{Q}_t^n$ and its yield is $\widehat{y}_t^n$ in this case.

2.2 Pricing kernels and currencies

Following B/F/T, we use affine no-arbitrage term structure models as a tool for investigating the relationship between bonds and currencies. We emphasize the distinction between the true, under the null of a model, pricing kernel $M_{t,t+i}$ and its estimate $M_{t,t+i}(\mathcal{D})$. Our notation $M_{t,t+i}(\mathcal{D})$ highlights that the estimated kernel is a function of the data $\mathcal{D}$ that is used in estimation. Specifically, we use various combinations of data on bonds $Q = \{Q_t^n\}$, $\widehat{Q} = \{\widehat{Q}_t^n\}$, and the corresponding exchange rate $S = \{S_t\}$. Thus, $\mathcal{D}$ is equal to a subset of $(Q, \widehat{Q}, S)$.

A long-standing tradition in the reduced-form no-arbitrage literature is to specify dynamics of the pricing kernel $M_{t,t+i}$ expressed in USD and that same kernel expressed in foreign currency, $\widehat{M}_{t,t+i}$. The latter implies a value of a foreign-currency-denominated foreign-issued bond

$$\widehat{Q}_t^n = E_t(\widehat{M}_{t,t+n}).$$

Estimating a model of $M_{t,t+1}$ and $\widehat{M}_{t,t+1}$ using data on Q and $\widehat{Q}$, one obtains estimates of the pricing kernels $M_{t,t+1}(Q)$, and $\widehat{M}_{t,t+1}(\widehat{Q})$.

Next, researchers infer the depreciation rate via

$$S_{t+1}/S_t = \widehat{M}_{t,t+1}(\widehat{Q})/M_{t,t+1}(Q). \tag{1}$$

There are variations in this approach where $M_{t,t+1}$ and $\widehat{M}_{t,t+1}$ are estimated simultaneously resulting in $M_{t,t+1}(Q, \widehat{Q})$ and $\widehat{M}_{t,t+1}(Q, \widehat{Q})$. Examples include, but are not limited to, Ahn (2004); Backus, Foresi, and Telmer (2001); Brennan and Xia (2006); Jotikasthira, Le, and Lundblad (2015); Kaminska, Meldrum, and Smith (2013). Sarno, Schneider, and Wagner (2012) are similar, but they also incorporate information about conditional expectations of the depreciation rates into their estimation procedure.

Most papers report that the depreciation rates from (1) do not resemble the observed depreciation rates. Most prominently, researchers document the forward premium and volatility anomalies. These results might simply manifest a model misspecification. However, the inherent empirical flexibility of affine models and sophistication of the authors involved suggest to us that bonds, on their own do not posses the information needed to capture the behavior of exchange rates.

2.3 Complete markets, hidden factors, and spanning

In this paper we argue that, in order to understand the described evidence, one needs to incorporate exchange rate data into a model of yield curves. That is because exchange rates are hidden in the yield curve. The differences between market completeness and hidden factors plays an important role in this paper. Both concepts are often associated with spanning. Thus, we define them here for clarity.

A market is complete with respect to a given set of assets if one can trade these assets to achieve any possible payoff and, therefore, know its value. We focus on whether markets are complete with respect to bonds alone, and we do not take a stand on the market completeness with respect to a larger set of assets.

This discussion is pertinent for the justification of the relationship (1). It is valid under two assumptions. First, markets are complete. Second, the estimated pricing kernel matches the true pricing kernel up to estimation noise, that is, Q and $\widehat{Q}$ span the space of all assets. Put differently, the markets are complete with respect to bonds.

A hidden factor arises in the context of multi-factor affine yield-curve models. The main property of a hidden factor is that it does not affect the cross-section of bond yields, yet it affects the joint dynamics of the factors that yields depend on. See Duffee (2011). A hidden factor cannot be replicated by a linear combination of yields. If the depreciation rate is a hidden factor then the markets are incomplete with respect to bonds.

The hidden-factor concept is different from market completeness. If the market is complete with respect to bonds and currencies, and depreciation rates are not hidden in yields, then the market is also complete with respect to the bonds only. However, if depreciation rates are hidden then the market could still be complete, just with respect to a larger set of assets. Further, markets may be incomplete regardless of the bonds' relation to exchange rates.

A popular empirical approach for establishing market completeness is to implement spanning regressions. One can think of regressing returns (payoffs with prices equal to 1) of an asset of interest on returns (payoffs) of the set of assets, with respect to which a market is hypothesized to be complete. An $R^2 < 1$ implies that the market is unconditionally incomplete (there is no static trading strategy that replicated a given payoff). The term structure literature implements regressions, refereed to as spanning also, of a candidate hidden factor on yields. To be clear, these are not regressions of payoffs, but, as we show subsequently, they have implications for payoffs.

3 The relation between exchange rates and bonds

The main purpose of this section is to provide motivating evidence about bonds' inability to span exchange rates. We do so both empirically and theoretically. First, we implement two types of regressions, the ones associated with market completeness and with hidden factors. Second, we offer a fully-worked out analytical example of a two-factor model with unspanned exchange rates. The purpose of this illustration is to construct how the true pricing kernel would look like in this setting, and how the estimated exchange rate would differ from the actual one.

3.1 Data

We work with monthly data from the US, UK, and Germany/Eurozone from January 1983 to December 2015 making for $T = 396$ observations per country. All data is aligned to the end of the month. US government yields are downloaded from the Federal Reserve and are constructed by Gurkaynak, Sack, and Wright (2007). All foreign government zero-coupon yields with maturities $12, 24, 36, 48$, and 60 months are downloaded from their respective central banks (Federal Reserve, Bank of England, and Bundesbank). The corresponding nominal exchange rates are from the Federal Reserve Bank of St. Louis. Prior to the introduction of the Euro, we use the German Deutschemark and splice these series together beginning in 1999.

Additionally, we use data on one-month yields to connect our approach to the evidence on UIP regressions. US one-month yields are downloaded from CRSP. UK and German yields are harder to obtain. We get two data sources for each (investing.com for both; Bank of England for the UK; Federal Reserve Bank of St. Louis for Germany) and use only data that match across the two sources. This approach produces a nearly full sample for the UK, and some missing observations for Germany from September 2007 through October 2010.

3.2 Principal components

We summarize properties of yields via principal component (PC) analysis. We consider three variations of the PC-construction. First, we extract six PCs from US bonds and the bonds of the country corresponding to the depreciation rates. Second, we extract six PCs from bonds of all three countries. Lastly, we extract nine PCs from bonds of all three countries. Table 1 reports the results. Six PCs explain 99.98% of variation in the yields of all three countries. Nine PCs across the three countries explain as much variation as six PCs do for yields of two countries (99.9995%).

Next, we approximate PCs with simpler linear combinations of yields to facilitate interpretation. Specifically, we introduce a vector

$$f_t = \begin{pmatrix} y_t^1 \\ \Delta_c^{€} y_t^1 \\ \Delta_c^{£} y_t^1 \\ y_t^{60,1} \\ \Delta_c^{€} y_t^{60,1} \\ \Delta_c^{£} y_t^{60,1} \end{pmatrix} = \begin{pmatrix} \text{US 1 month yield} \\ \text{US 1 month yield - Euro 1 month yield} \\ \text{US 1 month yield - UK 1 month yield} \\ \text{US slope = US 60 month yield - US 1 month yield} \\ \text{US slope - Euro slope} \\ \text{US slope - UK slope} \end{pmatrix}. \quad (2)$$

Here we use Δ to denote the one-period time-series difference operator, and Δ_c to denote the cross-country difference operator. Figure 1 shows that the PCs and these factors are similar.

3.3 Spanning regressions

Replication

One way to establish whether an asset (exchange rate) is spanned by other assets (bonds) is to regress returns of the former on the returns of the latter. To realize a return on an exchange rate, one must convert domestic currency into foreign currency, purchase a foreign (riskless) bond, sell it at a later date and then convert the proceeds back to the domestic currency. In order to avoid exposure to interest rate risk, this has to be a buy-and-hold strategy: $R_{t+1}^{FX} = S_{t+1}/S_t \times 1/\widehat{Q}_t^1$. A return on a domestic n-period bond is $R_{t+1}^n = Q_{t+1}^{n-1}/Q_t^n$. Thus, one can regress R_{t+1}^{FX} on a set of bond returns R_{t+1}^n for a variety of horizons n.

One practical problem with these regressions is that we do not have data on bonds with maturities that are one month apart to compute monthly returns. Thus, the regression could be implemented at an annual frequency only. We regress R_{t+12}^{FX} on R_{t+12}^n, $n = 12, 24, 36, 48, 60$. The R^2, regular and adjusted, from these regressions are reported in Table 2A, in the column labeled "\$ returns." The amount of variation in currency returns that can be hedged with bonds is quite modest, 16% at most.

It is tempting to take returns on foreign bonds, $\widehat{R}_{t+1}^n = \widehat{Q}_{t+1}^{n-1}/\widehat{Q}_t^n$, convert them to USD returns, $S_{t+1}/S_t \times \widehat{R}_{t+1}^n$, and add them on the right-hand side of the regression. But that would correspond to using exchange rates to span themselves.

Thus, to check if the exchange rate is spanned by foreign bonds, one must take the perspective of a foreign-currency investor: $\widehat{R}_{t+1}^{FX} = S_t/S_{t+1} \times 1/Q_t^1$, and regress these on foreign bond returns $\widehat{R}_{t+1}^n$. Table 2A, column labeled "€ or £ returns" reports the R^2, which are of the same magnitude as the USD-denominated returns. This evidence establishes that bonds are unable to span the space of currency payoffs.

That one can use only assets traded in a given country for replicating the corresponding FX rate suggests why the documented lack of spanning is a natural result. In fact, it would be unusual if one would be able to trade bonds of a given country in such a way that they trace out its FX rate. This point does not have to be about bonds alone. While clearly outside of the focus of this paper, we illustrate this by complementing the bond returns with MSCI stock index returns of the corresponding countries in the last two columns of Table 2A. The maximal R^2 moves to 28%, which is still very far from successful spanning. The main asset classes are only weakly related to exchange rates.[1]

This result is also natural in the context of baseline international real business cycle models. In the case of complete financial markets, e.g., Backus, Kehoe, and Kydland (1994), an exchange rate can be replicated by the full set of Arrow-Debreu securities. So, unless the theoretical setting is such that bonds are capable of completing the financial markets, one would expect bonds to be unable to replicate exchange rates.

One strand of the literature argues for incomplete financial markets in order to be able to understand empirical evidence on the lack of consumption risk sharing. In such a setting, depending on dynamics of output shocks or how relative prices of imports and exports (terms of trade) react to these shocks, one could obtain a nearly perfect risk sharing (Baxter and Crucini, 1995, Cole and Obstfield, 1991), or lack of thereof (e.g., Corsetti, Dedola, and Leduc, 2008). Our evidence points towards the latter.

Hidden factors in affine models

We implement an additional set of regressions to understand how to model bonds and currencies in an affine context. Consequently, we focus on establishing whether the depreciation rate is a hidden factor in domestic and foreign yield curves, i.e., that it does not appear as a factor in the cross-section (Duffee, 2011). As we demonstrate in the next subsection, a variable that is hidden in the yield curve is not spanned dynamically by a portfolio of bonds.

There is a well-established literature that primarily focuses on US bonds and seeks to identify hidden variables. Joslin, Priebsch, and Singleton (2014) refer to them as unspanned and use regressions of inflation or output growth on yields to motivate their use in a term structure model. We follow these authors and implement a regression to motivate our conjecture that depreciation rates are unspanned by yields. We regress the individual log depreciation rate on the three versions of principal components (PC) constructed from yields discussed in section 3.2. The timing is $\Delta s_t \sim f_t$ at a monthly frequency.

[1] One can think of more creative equity portfolios that could have a stronger relation to exchange rates, e.g., global firms, or commodity-intensive industries. Going down this path would lead to an investigation of market completeness, which is not the point of our paper. Our objective here is to illustrate that it is natural that some "typical" assets have low correlation with exchange rates.

Table 2B displays the regression results. The largest R^2 is 6.38%.[2] Depreciation rates have very little to do with bond yields. To be clear, we are not imposing this conclusion on our main term structure model that we introduce later in the paper. This conclusion serves as a motivation for considering different versions of that model. We will be formally testing, in the context of our model, whether depreciation rates are unspanned.

4 Implications of unspanned exchange rates

This section discusses the relationship between two literatures, (i) exchange rate modeling using pricing kernels, and (ii) international yield curve modeling. Most papers treat these undertakings separately despite the common pricing kernel approach and affine modeling tools. Historically, one group of researchers targeted exchange rate moments and did not use yields in estimation. A subgroup of these researchers found that a fit to yields is bad. Another group of researchers targeted yield moments and did not use exchange rates in estimation. A subgroup of those researchers evaluated the implications for the moments of exchange rates and found that their implied dynamics are very far from reality.

We emphasize commonality between bond and exchange rate modeling. We offer a way to unify all the evidence by accounting for the fact that exchange rates are not spanned by bonds. A simple implication of this observation is that one has to use exchange rates jointly with bond prices to estimate models of pricing kernels. In the next section, we demonstrate that this approach leads to a realistic model of both bonds and exchange rates.

4.1 The asset market view of exchange rates

Revisiting equation (1), we re-write it in logs and in two steps:

$$\Delta s_{t+1} = \widehat{m}_{t,t+1} - m_{t,t+1} = \widehat{m}_{t,t+1}(Q, \widehat{Q}) - m_{t,t+1}(Q, \widehat{Q}).$$

The first equality reflects the common use of the asset market view in the literature. The second equality emphasizes the role of data used to estimate the pricing kernels.

The relationship holds under two assumptions. First, the markets are complete, which gives the first equality. This is well understood and if often assumed explicitly. The second assumption is that the estimated pricing kernel matches the true pricing kernel up to estimation noise, that is, Q and $\widehat{Q}$ span the space of all assets. We presented evidence in Section 3.3 that this assumption is violated.

[2] As a reference, Joslin, Priebsch, and Singleton (2014), who argue that inflation and output growth are unspanned by bonds, report R^2 values of 86% and 32%, respectively.

4.2 Cross-country differences between bond yields

The international bond literature focuses on the strong common factor structure across US and foreign yields. It leads to a conclusion that global (a.k.a., U.S.) variables are responsible for most of the yield variation. We argue that there are economically important cross-country differences between yields and bond risk premiums. These differences are driven by the properties of exchange rates.

The yields y_t^n and $\widehat{y}_t^n$ share a simple relationship. Under conditional log-normality,

$$\begin{aligned}\widehat{q}_t^n &= \log E_t\left(e^{\sum_{i=1}^n m_{t+i-1,t+i}+\Delta s_{t+i}}\right)\\ &= q_t^n + E_t\left(\sum_{i=1}^n \Delta s_{t+i}\right) + \frac{1}{2}var_t\left(\sum_{i=1}^n \Delta s_{t+i}\right) + cov_t\left(\sum_{i=1}^n m_{t+i-1,t+i}, \sum_{i=1}^n \Delta s_{t+i}\right).\end{aligned}$$

After multiplying both sides by $-n^{-1}$, we get the interest rate differential (IRD):

$$\Delta_c y_t^n \equiv y_t^n - \widehat{y}_t^n = es_t^n - srp_t^n + vs_t^n. \tag{3}$$

Here $es_t^n \equiv n^{-1}E_t\left(\sum_{i=1}^n \Delta s_{t+i}\right)$ is the average expected depreciation rate, and $srp_t^n \equiv -n^{-1}cov_t\left(\sum_{i=1}^n m_{t+i-1,t+i}, \sum_{i=1}^n \Delta s_{t+i}\right)$ is the (ex-ante) currency "risk premium." We use the quotation marks because srp_t^n does not reflect the convexity term $vs_t^n \equiv (2n)^{-1}var_t\left(\sum_{i=1}^n \Delta s_{t+i}\right)$. The currency risk premium measures the additional compensation that an investor in foreign bonds requires in order to be exposed to future shocks to the exchange rate.

The combination of the last column of Table 1 and Figure 1 tells us that the short-term IRDs, $\Delta_c y_t^1$, serve as PC2 and PC3, approximately. Because of the connection of this spread to expected depreciation rates via equation (3), it is immediately clear that exchange rates play an important role in the behavior of international yield curves. While the US level obviously plays the major role by explaining 93.5% of variation, PC2 and PC3 contribute 4.4% and 1.4%, respectively. These are non-trivial amounts as is well known from the literature on US-only yield curve modeling.

Also, there is a simple currency-related connection between the excess returns on bonds from different countries. The USD bond one-period excess returns, in logs, are:

$$rx_{t+1}^n \equiv q_{t+1}^{n-1} - q_t^n + q_t^1 = -(n-1)y_{t+1}^{n-1} + ny_t^n - y_t^1 \tag{4}$$

with a similar expression for the foreign currency, $\widehat{rx}_{t+1}^n$. Note that the reference rate for foreign excess returns is the short rate of the respective country, $\widehat{y}_t^1$. Therefore, $\widehat{rx}_{t+1}^n$ does not depend on the currency of that country. In logs, this is equivalent to using the US short rate y_t^1 as a reference irrespective of the country and then constructing currency-hedged bond returns.

Combining equations (4) and (3) we get:

$$\begin{aligned}\Delta_c rx^n_{t+1} \equiv rx^n_{t+1} - \widehat{rx}^n_{t+1} &= (n-1)\cdot srp^{n-1}_{t+1} - n\cdot srp^n_t + srp^1_t \\ &- (n-1)\cdot vs^{n-1}_{t+1} + n\cdot vs^n_t - vs^1_t \\ &- u^n_{t+1} + u^1_{t+1},\end{aligned} \tag{5}$$

where, for a given horizon j, $u^j_{t+1} = E_{t+1}\left(\sum_{i=1}^j \Delta s_{t+i}\right) - E_t\left(\sum_{i=1}^j \Delta s_{t+i}\right)$ – is the surprise in expectations of the depreciation rate. Therefore, ignoring convexity, differences in expected log excess returns are driven by the differences in currency risk premiums across different horizons.

Equations (3) and (5) demonstrate that researchers studying bond yields and bond risk premiums are inherently studying the same object as researchers studying the dynamics of currency risk premiums. The twist is that it would be natural to think that factors driving one type of risk premium should show up as factors driving the other. However, that is not the case because exchange rates are unspanned by bonds. There is at least one factor that affects currency premiums but does not appear in bond premiums.

Therefore, the observed differences between bond yields or risk premiums, on their own, do not allow identifying the currency risk premiums, srp^n_t, even if the variance of the depreciation rate is constant. In the case of the yield differential, the premium is "contaminated" by the expected depreciation rate. In the case of the bond premium differential, the premium is affected by the different timing and horizons. One needs an explicit model of the depreciation rate that accounts for the lack of spanning to disentangle the currency premium and other components.

4.3 Illustration of unspanned exchange rates

The presented evidence implies that, depending on whether one uses bond data alone $(Q, \widehat{Q})$, or combines bond data with exchange rates S for estimation, the pricing kernel may have very different properties. That is the point of departure from the exploration of incomplete markets in affine settings studied by B/F/T and Lustig and Verdelhan (2015). These authors specify a model of the true domestic and foreign pricing kernels, $M_{t,t+1}$ and $\widehat{M}_{t,t+1}$, respectively. Then they discuss conditions under which the markets could be incomplete and how many assets are required to span the markets. Despite similar-looking equations, we do not assert dynamics of the full pricing kernel. Instead, we specify the dynamics of that part of the kernel that matches the properties of a given set of assets: $(Q, \widehat{Q})$ in the NFX case; $(Q, \widehat{Q}, S)$ in the WFX case.[3]

[3] Sandulescu, Trojani, and Vedolin (2018) establish conditions under which the ratio of the two pricing kernels recovers the exchange rate correctly even if markets are incomplete. The first condition is that domestic and foreign markets are integrated, that is, the span of domestic asset returns coincides with the span of foreign asset returns after conversion into the domestic currency. Our spanning regression is a

Before proceeding with the estimation, we introduce a simple model of the true pricing pricing kernel in order to demonstrate the effects of hidden depreciation rates in theory and in practice, when these rates are not used for estimation. We emphasize four main points.

First, even if markets are complete with respect to exchange rates and bonds, and one knows the true model of the pricing kernel, the estimated depreciation rate cannot be identified if it is not observed. As a result, the estimated expected depreciation rate and its volatility are biased. Second, because of that, one cannot realistically decompose yield or bond premium spreads into the expected depreciation rate and currency risk premium. Third, the modeling of bond yields is unaffected by unobserved exchange rates. Fourth, once the exchange rate is introduced into the set of observations, one does not need market completeness to estimate its dynamics accurately.

Setup

The model can be described as the international version of the Vasicek (1977) term structure model. Consider a vector $x_t = (y_t^1, \Delta s_t)^\top$ that follows a VAR(1):

$$x_t \quad = \quad \mu_x + \Phi_x x_{t-1} + \Sigma_x \varepsilon_t \qquad\qquad \varepsilon_t \sim \mathrm{N}\left(0, I\right), \tag{6}$$

where the vector μ_x has elements $\mu_{x,i}$, matrix Φ_x has elements $\phi_{x,ij}$, and matrix Σ_x has elements $\sigma_{x,ij}$.

We model the dynamics of the true log pricing kernel expressed in USD

$$m_{t,t+1} \quad = \quad -y_t^1 - \frac{1}{2}\lambda_t^\top \lambda_t - \lambda_t^\top \varepsilon_{t+1}, \tag{7}$$

with market prices of risk

$$\lambda_t \quad = \quad \Sigma_x^{-1}\left(\lambda_0 + \lambda_x x_t\right). \tag{8}$$

For the purposes of this section only, we assume that the matrix λ_x has a special form:

$$\lambda_x = \begin{pmatrix} \lambda_{x,11} & \phi_{x,12} \\ \lambda_{x,21} & \phi_{x,22} \end{pmatrix}.$$

In words, changes in interest rate and currency risk premiums due to variations in the depreciation rate exactly offset changes in expectations of future short interest rates and future depreciation rates, respectively.

particular case of that when there is only one foreign asset return, that is, the return on a one-period bond. Because we reject spanning, we reject integration (with respect to bonds) as well. The second condition, in a model-free setting, is that the estimated pricing kernels are the minimal entropy ones. We use affine models to connect to research on international yield curves.

Bond yields

The prices of zero-coupon USD-denominated bonds with maturity n are given by the standard pricing condition

$$Q_t^n = E_t\left(M_{t,t+1}Q_{t+1}^{n-1}\right).$$

As a result, because $\lambda_{x,12} = \phi_{x,12}$, US yields are linear functions of y_t^1 only

$$y_t^n = a_n + b_n y_t^1. \tag{9}$$

See Duffee (2011).

Let $\mathbf{e}_j$ denote a unit vector with a one in location j and zeros in all other entries. The currency risk premium is

$$srp_t^1 = -cov_t(m_{t,t+1}, \Delta s_{t+1}) = \mathbf{e}_2^\top \Sigma_x \lambda_t = \lambda_{0,2} + \lambda_{x,21} y_t^1 + \phi_{x,22}\Delta s_t. \tag{10}$$

We can express the pricing kernel in foreign currency:

$$\begin{aligned}\widehat{m}_{t,t+1} &= m_{t,t+1} + \Delta s_{t+1} = m_{t,t+1} + \mathbf{e}_2^\top x_{t+1} \\ &= -\widehat{y}_t^1 - \frac{1}{2}\widehat{\lambda}_t^\top \widehat{\lambda}_t - \widehat{\lambda}_t^\top \varepsilon_{t+1},\end{aligned} \tag{11}$$

where

$$\widehat{y}_t^1 = \alpha + \beta y_t^1, \tag{12}$$
$$\alpha = -\mathbf{e}_2^\top(\mu_x - \lambda_0) - \mathbf{e}_2^\top \Sigma_x \Sigma_x^\top \mathbf{e}_2/2 = -\mu_{x,2} + \lambda_{0,2} - (\sigma_{x,21}^2 + \sigma_{x,22}^2)/2, \tag{13}$$
$$\beta = 1 - \phi_{x,21} + \lambda_{x,21}, \tag{14}$$
$$\widehat{\lambda}_t = \Sigma_x^{-1}\left(\widehat{\lambda}_0 + \lambda_x x_t\right), \tag{15}$$
$$\widehat{\lambda}_0 = \lambda_0 - \Sigma_x \Sigma_x^\top \mathbf{e}_2. \tag{16}$$

The foreign interest rate $\widehat{y}_t^1$ in (12) does not depend on the depreciation rate because $\lambda_{x,22} = \phi_{x,22}$. The foreign-currency denominated U.S. pricing kernel $\widehat{m}_{t,t+1}$ in (11) is equal to the true foreign pricing kernel if markets are complete but may not be otherwise.

The prices of zero-coupon foreign currency bonds with maturity n are given by

$$\widehat{Q}_t^n = E_t\left[\widehat{M}_{t,t+1}\widehat{Q}_{t+1}^{n-1}\right].$$

As a result, because $\lambda_{x,22} = \phi_{x,22}$, foreign yields are linear functions of $\widehat{y}_t^1$ only

$$\widehat{y}_t^n = \widehat{a}_n + \widehat{b}_n \widehat{y}_t^1. \tag{17}$$

In particular, it implies that $\Delta_c y_t^n$ and $E_t \Delta_c r x_{t+1}^n$ do not depend on Δs_t.

Spanning

Equations (9) and (17) imply that the depreciation rate is hidden in the yield curves. We can characterize the implication of these relationships for spanning of currency returns with a portfolio of bonds that was described in section 3.3:

$$
\begin{aligned}
R^{FX}_{t+1} &= e^{\Delta s_{t+1}+\widehat{y}^1_t} \text{ spanned with } R^n_{t+1} = e^{a_n - a_{n-1} + b_n y^1_t - b_{n-1} y^1_{t+1}};\\
\widehat{R}^{FX}_{t+1} &= e^{-\Delta s_{t+1}+y^1_t} \text{ spanned with } \widehat{R}^n_{t+1} = e^{\widehat{a}_n - \widehat{a}_{n-1} + \widehat{b}_n \widehat{y}^1_t - \widehat{b}_{n-1} \widehat{y}^1_{t+1}}.
\end{aligned}
$$

The analytics simplify if we express the return on a portfolio of bonds with a vector of weights w_t in logs. The switch to logs can be justified via a log-linearization: $\log(\tilde{w}_t^\top e^{r_{t+1}}) \approx w_{t0} + w_t^\top r_{t+1}$. We use additional notation: b $(\widehat{b})$ for a vector of loadings b_{n-1} $(\widehat{b}_{n-1})$ for a range of maturities n, and r_{t+1} $(\widehat{r}_{t+1})$ for a vector of log returns $\log R^n_{t+1}$ $(\log \widehat{R}^n_{t+1})$. Then,

$$
\begin{aligned}
|corr_t(r^{FX}_{t+1}, w_t^\top r_{t+1})| &= \frac{cov_t(\Delta s_{t+1}, w_t^\top b y^1_{t+1})}{var_t^{1/2}(\Delta s_{t+1}) var_t^{1/2}(w_t^\top b y^1_{t+1})} = \frac{\sigma_{x,21}}{(\sigma^2_{x,21}+\sigma^2_{x,22})^{1/2}} < 1,\\
|corr_t(\widehat{r}^{FX}_{t+1}, w_t^\top \widehat{r}_{t+1})| &= \frac{cov_t(\Delta s_{t+1}, w_t^\top \widehat{b} \widehat{y}^1_{t+1})}{var_t^{1/2}(\Delta s_{t+1}) var_t^{1/2}(w_t^\top \widehat{b} \widehat{y}^1_{t+1})} = \frac{\sigma_{x,21}}{(\sigma^2_{x,21}+\sigma^2_{x,22})^{1/2}} < 1.
\end{aligned}
$$

These correlations imply that $R^2 < 1$ for the spanning regressions.[4]

Thus, a conditional portfolio of bonds cannot span an exchange rate if it is hidden in the yield curve. Nevertheless, if we assume existence of the currency market and the ability to trade an infinite number of bonds, the markets would be complete. In that case $\widehat{m}_{t,t+1}$ would be the true foreign pricing kernel.

Estimation, NFX approach

The NFX approach writes down separate pricing kernels for the foreign and domestic countries. It then ignores information in the depreciation rates to estimate the term structure model. A natural question to ask is whether we can recover the true dynamics of Δs_t by using bonds alone. In this thought experiment, we ignore estimation uncertainty by assuming that we have an infinite amount of data on foreign and domestic bonds. The true dynamics of Δs_t are:

$$\Delta s_{t+1} = \mu_{x,2} + \phi_{x,21} y^1_t + \phi_{x,22}\Delta s_t + \sigma_{x,21}\varepsilon_{1t+1} + \sigma_{x,22}\varepsilon_{2t+1}. \tag{18}$$

Suppose that, in addition, depreciation rates do not predict interest rates (consistent with the evidence to be discussed in a later section), that is, $\phi_{x,12} = 0$, and Σ_x is lower triangular

[4] If the depreciation rate were not hidden, then the conditional correlations above would depend on the portfolio weights. These weights then could be chosen to maximize the correlation with a possibility of reaching the value of 1 (if markets are complete with respect to bonds and currencies).

(recursive identification). Then equation (6) implies that the interest rate y_t^1 follows a process

$$y_t^1 = \mu_{x,1} + \phi_{x,11} y_{t-1}^1 + \sigma_{x,11}\varepsilon_{1t},$$

which can be easily estimated. In addition to the parameters in this equation, the coefficients a_n and b_n in equation (9) are non-linear functions of $\lambda_{0,1}$ and $\lambda_{x,11}$. These parameters are identified from US bond data and can be estimated using non-linear least squares.

The estimated USD pricing kernel is

$$m_{t,t+1}(Q,\widehat{Q}) = m_{t,t+1}(Q) = -y_t^1 - [\sigma_{x,11}^{-1}(\lambda_{0,1} + \lambda_{x,11}y_t^1)]^2/2 - \sigma_{x,11}^{-1}(\lambda_{0,1} + \lambda_{x,11}y_t^1)\varepsilon_{1t+1}$$

where all of the parameters can be recovered from linear or non-linear least squares regressions.

The foreign short rate $\widehat{y}_t^1$ is a linear transformation of y_t^1 in equation (12). If we do not use Δs_t in estimation, we can still identify the parameters α and β in (12) from a linear regression. The coefficients $\widehat{a}_n$ and $\widehat{b}_n$ in equation (17) are non-linear functions of $\widehat{\lambda}_{0,1}$ and $\lambda_{x,11}$. These parameters are identified from foreign bonds and can be estimated by non-linear least squares. Equation (16) implies that

$$\widehat{\lambda}_{0,1} = \lambda_{0,1} - \sigma_{x,11}\sigma_{x,22},$$

Given that we can identify $\lambda_{0,1}$ and $\sigma_{x,11}$ from US bonds, we can infer $\sigma_{x,22}$ from this equation.

The estimated foreign currency pricing kernel is

$$\widehat{m}_{t,t+1}(Q,\widehat{Q}) = -\widehat{y}_t^1 - [\sigma_{x,11}^{-1}(\widehat{\lambda}_{0,1} + \lambda_{x,11}y_t^1)]^2/2 - \sigma_{x,11}^{-1}(\widehat{\lambda}_{0,1} + \lambda_{x,11}y_t^1)\varepsilon_{1t+1}.$$

with all of the parameters identified from linear or non-linear regressions.

Thus, if a researcher were to use the complete markets logic to infer the depreciation rate, they would obtain

$$\begin{aligned}
\Delta s_{t+1}(Q,\widehat{Q}) &= \widehat{m}_{t,t+1}(Q,\widehat{Q}) - m_{t,t+1}(Q,\widehat{Q}) \\
&= \mu_{x,2} - \lambda_{0,2} + \sigma_{x,11}^{-1}\sigma_{x,22}\lambda_{0,1} + \sigma_{x,21}^2/2 \\
&+ (\phi_{x,21} - \lambda_{x,21} + \sigma_{x,11}^{-1}\sigma_{x,22}\lambda_{x,11})y_t^1 + \sigma_{x,22}\varepsilon_{1t+1}.
\end{aligned}$$

where have used the relationships in (13)-(16) to simplify these expressions. The implied currency risk premium is

$$srp_t^1(Q,\widehat{Q}) = -cov_t(m_{t,t+1}(Q,\widehat{Q}), \Delta s_{t+1}(Q,\widehat{Q})) = \sigma_{x,11}^{-1}\sigma_{x,22}\lambda_{0,1} + \sigma_{x,11}^{-1}\sigma_{x,22}\lambda_{x,11}y_t^1.$$

Comparing these expressions with the true depreciation rate in (18) and the true currency risk premium in (10) we see that we get a bias in the expected depreciation rate, risk premium (forward premium anomaly), and volatility (volatility anomaly).

Relatedly, Ahn (2004) constructs the depreciation rate to be a hidden factor by allowing Δs_t to be affected by a shock that does not affect the USD pricing kernel $m_{t,t+1}$. If we set the risk premium parameters $\lambda_{0,2}$, $\lambda_{x,21}$, $\lambda_{x,12}$, $\lambda_{x,22}$ and persistence parameters $\phi_{x,12}$, $\phi_{x,22}$ to zero, we obtain that model. Then the NFX method recovers the USD pricing kernel, but not the depreciation rate.

Estimation, WFX approach

The WFX approach uses observations on Δs_t for estimation. As a result one can estimate the full dynamics of x_t in (6). The transition from $m_{t,t+1}(Q, \widehat{Q}, S)$ to $\widehat{m}_{t,t+1}(Q, \widehat{Q}, S)$ by changing denomination via $\widehat{m}_{t,t+1}(Q, \widehat{Q}, S) = m_{t,t+1}(Q, \widehat{Q}, S) + \Delta s_{t+1}$ does not require any assumptions because it is a simple change in the denomination of the pricing kernel. Once the parameters of Δs_t are estimated, one can use α in (13) to back out the risk premium parameter $\lambda_{0,2}$ and then estimate β in (14) to back out $\lambda_{x,21}$.

Implications for international yield curves and currencies

If the depreciation rate is unspanned by bonds, then both estimation approaches would produce identical domestic and foreign yields curves. The WFX approach would allow one to use equations (3) and (5) to decompose the cross-country differences in yields into the currency risk premium and expected currency components. Conversely, since the NFX approach cannot accurately recover the dynamics of the depreciation rate, it cannot separately identify these components.

From the currency-modeling perspective, the WFX approach allows for identification of realistic dynamics of the depreciation rate. It also allows researchers to connect currency risk premiums to factors driving bond premiums. That is not feasible if one uses observations on currencies alone.

5 The full model

In this section we specify our general multi-factor model. Here we take a view that we do not know the true pricing kernel. Instead, we estimate the specified functional form using different datasets ($Q, \widehat{Q}$ for the NFX approach, and $Q, \widehat{Q}, S$ for the WFX approach). Further, we do not make any assumptions about the spanning of exchange rates. We test for that upon the model estimation. We model the dynamics of the state vector x_t as a Gaussian VAR given by (6). We use $\bar{\mu}_x$ to denote the unconditional mean of the state.

It has been argued in the literature that one needs to incorporate time-varying volatility to resolve the FX volatility anomaly in the context of the NFX approach (Anderson, Hammond, and Ramezani, 2010; Brandt and Santa-Clara, 2002). Of course, FX volatility is time-varying and adding that feature is an obvious extension if one is interested in FX option valuation or other aspects of FX dynamics. We do not model stochastic volatility to emphasize the point that there is no volatility anomaly even in a Gaussian model if the WFX approach is used.

Finally, B/F/T and many authors following them explore so-called square-root, or CIR, state variables instead of the Gaussian ones used here. This distinction plays no role here. We are simply looking for models that are capable of a realistic fit to the yield curve data. Starting from Dai and Singleton (2000) and many papers following them, the literature has concluded that Gaussian models are more flexible in capturing yield co-movement and risk premiums. These models have been a de-facto standard for the last 15 years. A square-root factor could be helpful in capturing time-varying volatility of interest rates, but, absent data on interest rate derivatives, it is very hard to identify empirically (Bikbov and Chernov, 2011).

5.1 Bonds denominated in US dollars

We assume the dynamics of the state x_t has the same functional form as the state in (6). We postpone selection of a specific state until a later section. We model the dynamics of the log pricing kernel expressed in USD

$$m_{t,t+1}(\mathcal{D}) \quad = \quad -\delta_{i,0} - \delta_{i,x}^{\top} x_t - \frac{1}{2}\lambda_t^{\top}\lambda_t - \lambda_t^{\top}\varepsilon_{t+1},$$

with market prices of risk having the same dependence on x_t as in equation (8). $\mathcal{D}$ highlights that we are not specifying the true pricing kernel $M_{t,t+1}$, but only its component that correctly prices a set of assets to be specified later, either $(Q, \widehat{Q})$, or $(Q, \widehat{Q}, S)$. Regardless of the choice of assets, our model of the component of the USD pricing kernel that values them correctly is the same.

The prices of zero-coupon USD-denominated bonds with maturity n are given by the standard pricing condition

$$Q_t^n \quad = \quad E_t\left(M_{t,t+1}(\mathcal{D})Q_{t+1}^{n-1}\right).$$

As a result, US yields are linear functions of the factors

$$y_t^n \quad = \quad a_n + b_{n,x}^{\top} x_t.$$

Expressions for the bond loadings can be found in Appendix A.1.

5.2 Bonds denominated in foreign currency

5.2.1 The NFX approach

In this case $M_{t,t+1}(\mathcal{D}) = M_{t,t+1}(Q, \widehat{Q})$. Here the state vector x_t is presumed to span all bonds. We model the dynamics of the log pricing kernel denominated in foreign currency as

$$\widehat{m}_{t,t+1}(Q, \widehat{Q}) = -\widehat{\delta}_{i,0} - \widehat{\delta}_{i,x}^{\top} x_t - \frac{1}{2}\widehat{\lambda}_t^{\top}\widehat{\lambda}_t - \widehat{\lambda}_t^{\top}\varepsilon_{t+1} \tag{19}$$

with market prices of risk

$$\widehat{\lambda}_t = \Sigma_x^{-1}\left(\widehat{\lambda}_0 + \widehat{\lambda}_x x_t\right).$$

We suppress country-specific notation for simplicity.

The prices of zero-coupon foreign currency bonds with maturity n are given by

$$\widehat{Q}_t^n = E_t\left[\widehat{M}_{t,t+1}(Q, \widehat{Q})\widehat{Q}_{t+1}^{n-1}\right].$$

As a result, foreign yields are linear functions of the factors

$$\widehat{y}_t^n = \widehat{a}_n^N + \widehat{b}_{n,x}^{N\top} x_t.$$

where the bond loadings can be found in Appendix A.2.

This strategy is similar to the ones undertaken in the literature on international yield curves. There is some variation: some authors distinguish between global and country-specific factors; some authors estimate the kernel expressed in USD using USD bond data only, $M_{t,t+1}(Q)$, as a first step, and then proceed with estimating $\widehat{M}_{t,t+1}(Q, \widehat{Q})$. We use the joint data and allow yields from each country to load on all the factors allowing the data to speak to which bonds load on which factors. Assuming market completeness with respect to bonds, one can infer the log depreciation rate as:

$$\Delta s_{t+1} = \widehat{m}_{t,t+1}(Q, \widehat{Q}) - m_{t,t+1}(Q, \widehat{Q}). \tag{20}$$

5.2.2 The WFX approach

In this case $M_{t,t+1}(\mathcal{D}) = M_{t,t+1}(Q, \widehat{Q}, S)$. Here the state vector x_t is presumed to span all bonds and exchange rates. The log depreciation rate Δs_t is assumed to be a linear function of the state vector:

$$\Delta s_t = \delta_{s,0} + \delta_{s,x}^{\top} x_t. \tag{21}$$

The prices of zero-coupon foreign currency bonds with maturity n are given by

$$\widehat{Q}_t^n = E_t\left(M_{t,t+1}(Q,\widehat{Q},S)\frac{S_{t+1}}{S_t}\widehat{Q}_{t+1}^{n-1}\right). \tag{22}$$

As a result, foreign yields are linear functions of the factors

$$\widehat{y}_t^n = \widehat{a}_n^W + \widehat{b}_{n,x}^{W\top} x_t.$$

where the bond loadings can be found in Appendix A.3. Bauer and de los Rios (2014); Graveline and Joslin (2011) model the FX rate directly as well, but do not explore the implications of such an approach for the FX anomalies, or differences between the yield curves.

Given the estimated model, we can express the pricing kernel in foreign currency:

$$\widehat{m}_{t,t+1}(Q,\widehat{Q},S) = m_{t,t+1}(Q,\widehat{Q},S) + \Delta s_{t+1}. \tag{23}$$

This equation is an accounting identity that does not require any assumptions in contrast to (20). Combining equations (21) and (23), we can write

$$\widehat{m}_{t,t+1}(Q,\widehat{Q},S) = -\widehat{y}_t^1 - \frac{1}{2}\widehat{\lambda}_t^\top\widehat{\lambda}_t - \widehat{\lambda}_t^\top\varepsilon_{t+1},$$

where $\widehat{\lambda}_t$ has the same functional form as (15), and

$$\widehat{\lambda}_0 = \lambda_0 - \Sigma_x\Sigma_x^\top\delta_{s,x}. \tag{24}$$

Thus, in contrast to many theoretical models of exchange rates, $M(Q,\widehat{Q},S)$ and $\widehat{M}(Q,\widehat{Q},S)$ are asymmetric. The asymmetry arises via the constant component of the risk premium – an implication of the constant volatility model of depreciation rates. Changing that feature would introduce further asymmetry via the time-variation in risk premium.

6 Results

6.1 Empirical approach

We extend the estimation procedure of Joslin, Singleton, and Zhu (2011) to international yield curves and use Bayesian MCMC to implement it. The approach has two ingredients. First, risk premiums λ_t and $\widehat{\lambda}_t$ are estimated by specifying risk-adjusted dynamics of the state that is implied by the specification of the pricing kernels. The mapping into risk-adjusted parameters and identifying restrictions are discussed in Appendix B.1. Second, the state x_t is observable and is a linear transformation of yields and, in our case, exchange rates.

The choice of state vector is motivated by the PC analysis in section 3.2. Specifically, in the case of the NFX model, the state is $x_t^N = f_t$, see equation (2). In the WFX model, we complement the state vector by adding the two depreciation rates

$$x_t^W = \left(\Delta s_t^{€}, \Delta s_t^{£}, x_t^{N\top}\right)^\top.$$

Thus, our specification is closely related to that of Dumas and Solnik (1995) except for the choice of variables: one-period interest rates and lagged stock returns are selected to be their state variables.

Because all the state variables in x_t are observable, the free parameters that govern the dynamics of the state, $\mu_x, \Phi_x, \Sigma_x\Sigma_x^\top$, are identifiable directly from the VAR in equation (6). These parameters therefore require no identifying restrictions. Restrictions are required on the factor loadings and the risk-premium parameters λ_0 and λ_x. These restrictions are necessary to exactly identify the model and are not over-identifying restrictions.

6.2 Basic properties of the estimated models

6.2.1 Estimates and fit

We report the estimated parameters in Appendix C. Table 3 displays pricing errors. The overall message from this set of Tables is that both models fit the given collection of domestic and foreign yields similarly in terms of small errors and values of estimated parameters that are related to yields.

The differences in the approaches are manifested by two extra sets of risk premium parameters for the NFX model and by extra parameters corresponding to depreciation rates for the WFX model. The WFX model exhibits a slight deterioration in fitting yields at shorter maturities suggesting some tension in fitting yields and depreciation rates with the same set of state variables.

6.2.2 Exchange rates are unspanned by bonds

From the perspective of modeling the yield curve, the two depreciation rates are the new factors in the WFX model as compared to the NFX model. Figure 2 demonstrates that they are unspanned by showing how bonds of different countries load on the two depreciation rates in the WFX model. While there are some departures from zero, none of them are statistically significant. Given that the monthly standard deviation of each depreciation rate is about 0.029 (10% per year), the largest monthly movement in a yield (UK at 5 months) is 0.2 basis points ($0.7 \times 10^{-3} \times 0.029 \times 100^2$) for one standard deviation move in a depreciation rate. This is not an economically significant amount either.

In contrast to the literature on hidden/unspanned variables that focuses on their ability to forecast bond excess returns, we make no such claims here. Inspection of the estimated persistence matrix Φ_x in Table Appendix C.3 suggests that depreciation rates have little to do with forecasting of future yields.[5] Indeed, most of the elements in the first two columns of Φ_x are statistically close to zero.[6] There is one exception: the lagged GBP depreciation rate predicts the EUR one, and vice versa. Depreciation rates do not help forecasting bond risk premiums either as the relevant elements of the risk premium matrix λ_x are close to zero as well.

6.2.3 The role of other factors

We describe how the other factors impact yields by reporting the term structure of factor loadings in Figure 3. We show these loadings for the WFX model, while the NFX model (not reported) displays similar patterns. The figure should not be surprising. It shows that the US short interest rate and slope act as level and slope factors for all countries. The departures from the US factors, differences in short rates or differences in slopes, primarily affect yields of the corresponding foreign country.

These bond loadings could tempt a researcher to label US variables as global and the other variables as local, a language that is commonly used in the literature on international yield curve models. One has to be careful with such an interpretation because the choice of US factors as a de-facto reference is arbitrary. We could have considered an equivalent factor rotation where the German short rate and slope are the reference factors and the rest are defined relative to that.

Nevertheless, these loadings suggest a modular approach towards modeling multiple yield curves: start with a benchmark and its yield-based factors, then add as many countries as needed via factors that increment over the benchmark. One has to deal with parameter proliferation when more than three countries are considered at once. There is no choice but to impose overidentifying parameter restrictions; a problem we do not address in this paper. See Graveline and Joslin (2011) for a recent effort along these lines.

6.2.4 Implications for exchange rates

From the perspective of modeling depreciation rates, yields are an important ingredient in their dynamics. While risk-adjusted expected depreciation rates depend on their respective interest rate differential (IRD) only, an implication of covered interest parity (CIP), the

[5]Because of this property some of our readers have proposed to refer to exchange rates as irrelevant rather than hidden factors. We stick with the latter terminology because being hidden is the crucial feature one should have in an affine model to replicate the lack of spanning of exchange rates by bonds. Being irrelevant is an additional interesting property of exchange rates.

[6]That is consistent with our assumption $\phi_{x,12} = 0$ in section 4.3.

true expectations depend on other yield factors as well. At first glance, the model cannot replicate the violations of uncovered interest parity (UIP) documented by Bilson (1981); Fama (1984); Tryon (1979). For instance, the loading of the expected Euro depreciation rate on the respective IRD is positive. However, the estimated model is consistent with a multivariate regression with potentially correlated regressors, while the original UIP regressions are simple linear regressions with one regressor. In the next section, we explore what the WFX model implies for the specific univariate regressions studied in the literature.

While the WFX model matches the depreciation rate by construction, we need to infer them in the case of the NFX model. As pointed out in equation (20), we can do so only under the assumption of complete markets. Figure 4 displays the inferred and observed depreciation rate. They are clearly different in terms of their scale and dynamic patterns. In the next section we provide more details on these differences.

As noted in the context of equation (24), the modeled pricing kernels are asymmetric. The estimated model verifies that indeed estimated parameter values are such that asymmetry between $M(Q, \widehat{Q}, S)$ and $\widehat{M}(Q, \widehat{Q}, S)$ holds. Thus, if one were to build an equilibrium model with an objective to capture the evidence, the marginal rates of substitution of domestic and foreign economic agents might have to be asymmetric.

6.3 B/F/T

Results of B/F/T represent a challenge to affine no-arbitrage models. It appears to be difficult to replicate a certain set of stylized facts about interest rates and exchange rates simultaneously. They propose a model that succeeds in matching some of the properties of FX and yet generates unrealistic yield curves. Indeed, B/F/T state: "The implied yield curve ... is hump shaped with long yields reaching as high as 80 percent per annum." We revisit their analysis in the context of our models.

Table 4 replicates Table I of B/F/T in our sample and complements it by displaying model implications for the same set of facts. Both models do well in replicating facts about interest rates. This is not surprising given that B/F/T focused on short rates and they serve as state variables in our model. There is some deterioration for Euro/Germany because of the aforementioned missing observations on one-month yields.

The differences in FX implications are drastic. This is consistent with Figure 4, but offers other angles. The inferred depreciation rate is 25 times more volatile in the NFX model than in the data, and the mean is greater by two orders of magnitude. The NFX model is nowhere close to the results of UIP regressions. One way to interpret the NFX results is that markets are in fact incomplete with respect to bonds, so it is incorrect to use equation (20) to infer FX rates.

The WFX model replicates all of these moments perfectly, by construction. Does it come at the cost of a poor fit of the yield curve as in B/F/T? Table 3 foreshadows the answer.

Table 5 reports the yield-curve summary statistics in the B/F/T style. There is not much of a trade-off in fitting both yields and FX rates for the WFX model.

To clarify, while the B/F/T methodology is close to NFX, it is not identical. B/F/T construct their model to match the UIP violations and volatility of the depreciation rate. Thus, they use information about depreciation rates, but not about their joint dynamics with yields. That is why they can match some basic facts about currencies, but the resulting model cannot match yields. That is a manifestation of unspanned currencies. The only way for B/F/T to succeed empirically is to incorporate this lack of spanning into their model of exchange rates.

More generally, one might wonder if it is possible to use the presented evidence to construct M and $\widehat{M}$ such that they produce a realistic depreciation rate via (20). It is possible to do so, in theory, by taking the estimated $M(Q, \widehat{Q}, S)$ and $\widehat{M}(Q, \widehat{Q}, S)$. Section 4.3 offers a more explicit version of that in equations (7) and (11). The issue is, as pointed out in section 4.3, that one cannot estimate these models without the joint data on yields and currencies.

6.4 International yields and risk premiums

In this section we highlight implications of our analysis for international yield curve modeling. Our results suggest that if one were interested in fitting yields only, the NFX model would be sufficient for that task. However, a modest departure from this objective would render NFX incomplete.

Specifically, as equations (3) and (5) show, the cross-country differences between bond yields or risk premiums are solely driven by properties of currencies. So, if a researcher wanted to understand the sources of variation in the cross-country yield differentials, she would have to attribute that to the currency risk premium, srp_t^n, versus the other components. However, as we emphasize in section 4.2, neither the cross-country difference in bond yields nor risk premiums produces a direct measure of the currency risk premium for a given horizon.

The WFX model would help with that because it implies a realistic measure of srp_t^n by directly capturing the joint behavior of depreciation rates and prices of risk. Would it still be possible to decompose the yield differential, $\Delta_c y_t^n$, into the currency risk premium and expected depreciation rate using the NFX model? As we have shown, the only way to use the NFX model itself is to invoke equation (20) to infer an implicit depreciation rate and the corresponding risk premiums. That approach leads to a grossly misspecified depreciation rate.

An alternative approach would be to stick with the NFX model for its implications for yields. Next, a researcher could use an established approach towards currency premium measurement to produce an estimate of srp_t^n and use that to decompose $\Delta_c y_t^n$. We illustrate this strategy using the UIP regression to infer currency risk premiums.

6.4.1 Decomposition of cross-country differences in yields

We start with equation (3) and, following Campbell (1991), define news, at horizon n, about differences in yields, the expected depreciation rate, and currency risk premium

$$\begin{aligned} N^n_{\Delta y,t} &\equiv \Delta_c y^n_t - E_{t-1}\left(\Delta_c y^n_t\right), \\ N^n_{s,t} &\equiv es^n_t - E_{t-1}\left(es^n_t\right), \\ N^n_{srp,t} &\equiv srp^n_t - E_{t-1}\left(srp^n_t\right). \end{aligned}$$

Then, equation (3) implies

$$\begin{aligned} N^n_{\Delta y,t} &= N^n_{s,t} - N^n_{srp,t}, & (25) \\ var(N^n_{\Delta y,t}) &= var(N^n_{s,t}) + var(N^n_{srp,t}) - 2cov(N^n_{s,t}, N^n_{srp,t}). & (26) \end{aligned}$$

Thus, one can quantify the role of each component in the variation of the interest rate differential.

The same news components are affecting the decomposition of cross-country differences in bond risk premiums in equation (5). The difference in premiums can be re-written in terms of yields:

$$\Delta_c rx^n_t = -(n-1)\Delta_c y^{n-1}_t + n\Delta_c y^n_{t-1} - \Delta_c y^1_{t-1}.$$

Thus, news about the difference can be expressed in term of news about yields:

$$N^n_{\Delta rx,t} = \Delta_c rx^n_t - E_{t-1}\left(\Delta_c rx^n_t\right) = -(n-1)N^{n-1}_{\Delta y,t}.$$

In the sequel, we use the different approaches to estimate srp^n_t and the corresponding news to construct the decomposition (26). Anticipating the results, note that $var(N^n_{\Delta y,t})$ would be much smaller than any of the components on the right hand side simply because interest rates are, overall, much less variable than depreciation rates. So, for instance, reporting $var(N^n_{s,t})$ as a percentage of $var(N^n_{\Delta y,t})$ is not that informative.

Thus, we turn equation (25) around and work with:

$$N^n_{srp,t} = N^n_{s,t} - N^n_{\Delta y,t}. \quad (27)$$

The corresponding equation for variances is:

$$var(N^n_{srp,t}) = var(N^n_{s,t}) + var(N^n_{\Delta y,t}) - 2cov(N^n_{s,t}, N^n_{\Delta y,t}). \quad (28)$$

6.4.2 Decomposition results

We start with the WFX model, as a realistic benchmark. Given that the model is essentially a VAR, we compute all the relevant news in equation (27) directly from the model. The first row of Figure 5 displays the results.

Conveniently, the covariance term is close to zero. That is not surprising given the overall theme of a weak relationship between depreciation rates and interest rates. When $n = 1$ news about the cross-country yield differential has a very small contribution to the news about the currency premium, which is intuitive. As horizon grows the share of news about the cross-country yield differential starts growing and reaches about 50% at the 10-year horizon.

Next, as discussed above, we combine the NFX model with the UIP regression. The latter is used to establish a basic measure of currency risk premiums that is popular in the literature without taking a stand on spanning of currencies with bonds. Jointly, they form a restricted VAR. Thus, again, we can use standard techniques to construct news. The second row of Figure 5 displays the results.

The difference between the two approaches is striking. First, the contribution of both es_t and $\Delta_c y_t$ is greater than 100% at horizons below 6 months. That coincides with a positive covariance between the two components of the currency risk premiums. Second, even at longer horizons the pattern is qualitatively different from that of the WFX model. While the contribution of es_t is declining with horizon in the latter, it always stays above that of $\Delta_c y_t$. In the NFX/UIP case the contribution of es_t is always less, declining from about 80% at the 6-month horizon to 10% at the 10-year horizon. In contrast to the WFX model, the news about the yield differential contribute about 80% regardless of horizon.

7 Conclusion

We relate differences in international yield curves to exchange rates. In order to account for these risks, we combine estimation of yield curves with estimation of exchange rate dynamics. This exercise is unnecessary if bonds span exchange rates. We find drastic differences in results relative to a benchmark model estimated without exchange rates. Both models fit yields accurately, but the benchmark model implies exchange rates that are grossly incompatible with the observed data. Besides capturing realistic behavior of exchange rates, our main model speaks to the sources of the differences between the US and foreign yield curves, and their respective bond risk premiums.

Both differences are driven by currency risk premiums, and our model allows a researcher to decompose news about differences in bonds into news about currency risk premiums and expected depreciation rates. We show that a model that does not incorporate depreciation rates into estimation attributes too big of a contribution of news about the cross-country differences in yields to news about currency risk premiums.

References

Ahn, Dong-Hyun, 2004, Common factors and local factors: Implications for term structures and exchange rates, *Journal of Financial and Quantitative Analysis* 39, 69–102.

Anderson, Bing, Peter J. Hammond, and Cyrus A. Ramezani, 2010, Affine Models of the Joint Dynamics of Exchange Rates and Interest Rates, *Journal of Financial and Quantitative Analysis* 45, 1341–1365.

Backus, David, Silverio Foresi, and Chris Telmer, 2001, Affine term structure models and the forward premium anomaly, *Journal of Finance* 56, 279–304.

Backus, David K., Patrick J. Kehoe, and Finn E. Kydland, 1994, Dynamics of the trade balance and the terms of trade: The J-curve?, *American Economic Review* 84, 84–103.

Bauer, Gregory, and Antonio Diez de los Rios, 2014, An international dynamic term structure model with economic restrictions and unspanned risks, working paper, Bank of Canada.

Baxter, Marianne, and Mario Crucini, 1995, Business cycles and the asset structure of foreign trade, *Journal of International Economics* 36, 821–854.

Bikbov, Ruslan, and Mikhail Chernov, 2011, Yield curve and volatility: Lessons from eurodollar futures and options, *Journal of Financial Econometrics* 9, 66–105.

Bilson, John F. O., 1981, The speculative efficiency hypothesis., *Journal of Business* 54, 435–451.

Brandt, Michael, and Pedro Santa-Clara, 2002, Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets, *Journal of Financial Economics* 63, 161–210.

Brennan, Michael, and Yihong Xia, 2006, International Capital Markets and Foreign Exchange Risk, *The Review of Financial Studies* 19, 753–795.

Campbell, John Y., 1991, A variance decompostion for stock returns, *The Economic Journal* 101, 157–179.

Cole, Harold, and Maurice Obstfield, 1991, Commodity trade and international risk sharing: How much do finacial markets matter?, *Journal of Monetary Economics* 28, 3–24.

Corsetti, Giancarlo, Luca Dedola, and Sylvain Leduc, 2008, International risk sharing and the transmission of productivity shocks, *The Review of Economic Studies* 75, 443–473.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models., *The Journal of Finance* 55, 1943–1978.

Duffee, Gregory R., 2011, Information in (and not in) the term structure, *The Review of Financial Studies* 24, 2895–2934.

Dumas, Bernard, and Bruno Solnik, 1995, The world price of foreign exchange risk, *The Journal of Finance* 50, 445–479.

Fama, Eugene F., 1984, Forward and spot exchange rates., *Journal of Monetary Economics* 14, 319–338.

Graveline, Jeremy, and Scott Joslin, 2011, G10 swap and exchange rates, working paper.

Gurkaynak, Refet, Brian Sack, and Jonathan Wright, 2007, The U.S. Treasury yield curve: 1961 to the present, *Journal of Monetary Economics* 54, 2291–2304.

Joslin, Scott, Marcel Priebsch, and Kenneth J. Singleton, 2014, Risk premiums in dynamic term structure models with unspanned macro risks, *The Journal of Finance* 69, 1197–1233.

Joslin, Scott, Kenneth J. Singleton, and Haoxiang Zhu, 2011, A new perspective on Gaussian affine term structure models, *The Review of Financial Studies* 27, 926–970.

Jotikasthira, Chotibhak, Anh Le, and Christian Lundblad, 2015, Why do term structures in different currencies co-move?, *Journal of Financial Economics* 115, 58–83.

Kaminska, Iryna, Andrew Meldrum, and James Smith, 2013, A Global Model of International Yield Curves: No-Arbitrage Term Structure Approach, *International Journal of Finance and Economics* 18, 352–374.

Lustig, Hanno, and Adrien Verdelhan, 2015, Does incomplete spanning in international financial markets help to explain exchange rates?, working paper.

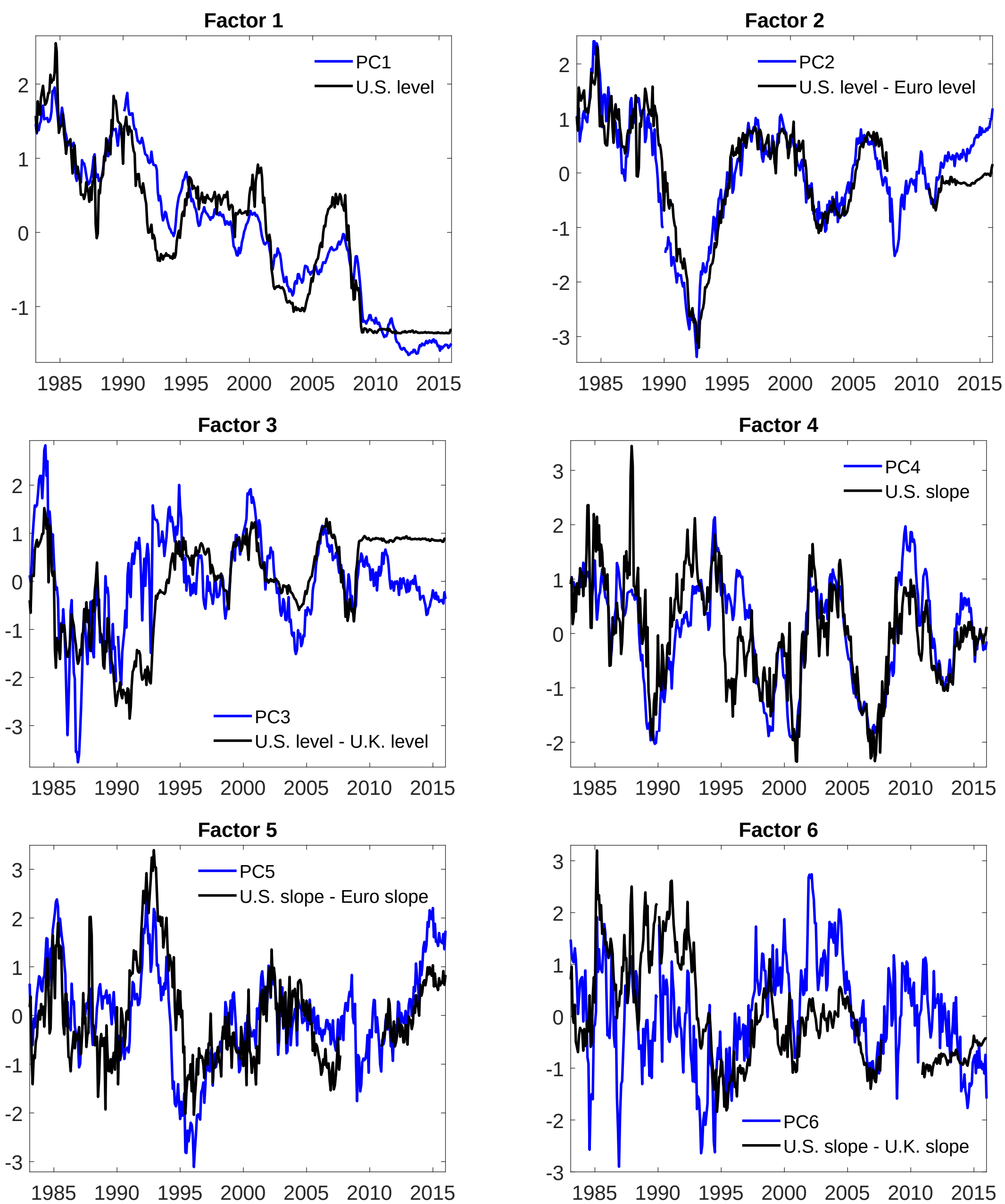
Sandulescu, Mirela, Fabio Trojani, and Andrea Vedolin, 2018, Model-free international stochastic discount factors, working paper.

Sarno, Lucio, Paul Schneider, and Christian Wagner, 2012, Properties of foreign exchange risk premiums, *Journal of Financial Economics* 105, 279–310.

Tryon, Ralph, 1979, Testing for rational expectations in foreign exchange markets, Federal Reserve Board international finance discussion paper no. 139.

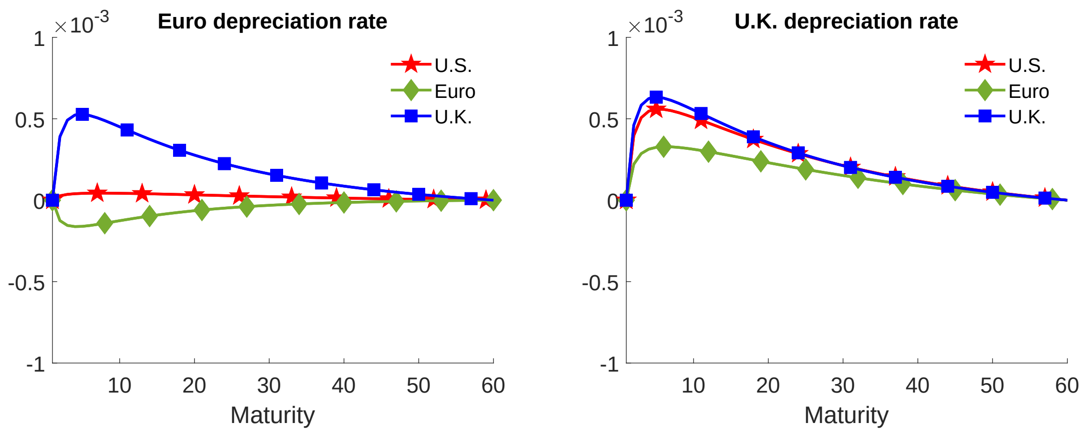
Vasicek, Oldrich Alfons, 1977, An equilibrium characterization of the term structure., *Journal of Financial Economics* 5, 177–188.

Figure 1
Principal components and yield factors



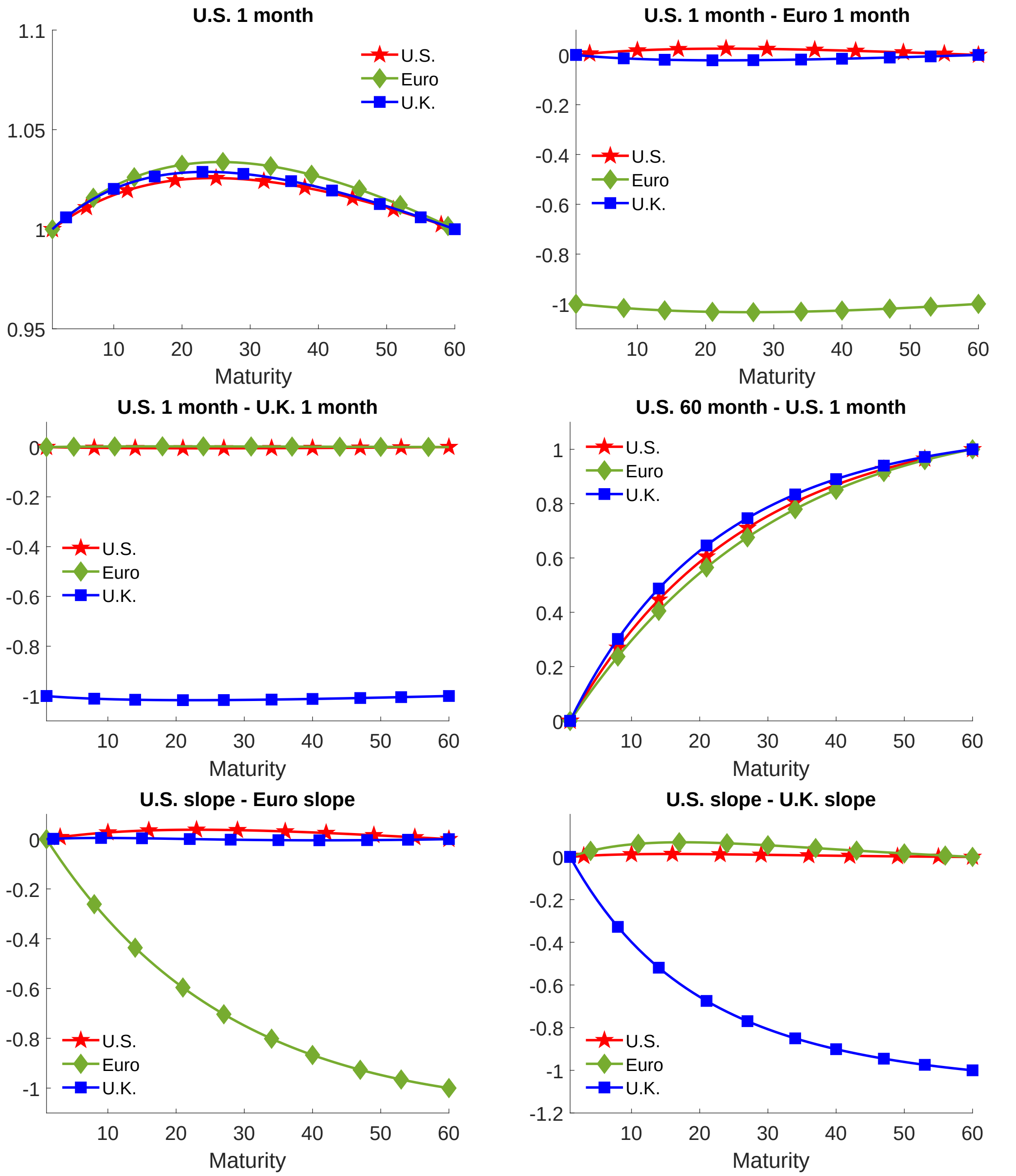
Notes: Plots of the first six principal components (blue) and the yield factors in the state vector x_t*. All variables have been standardized to have mean zero and variance one in the plot to make them comparable.*

Figure 2
Depreciation rate loadings for the WFX model



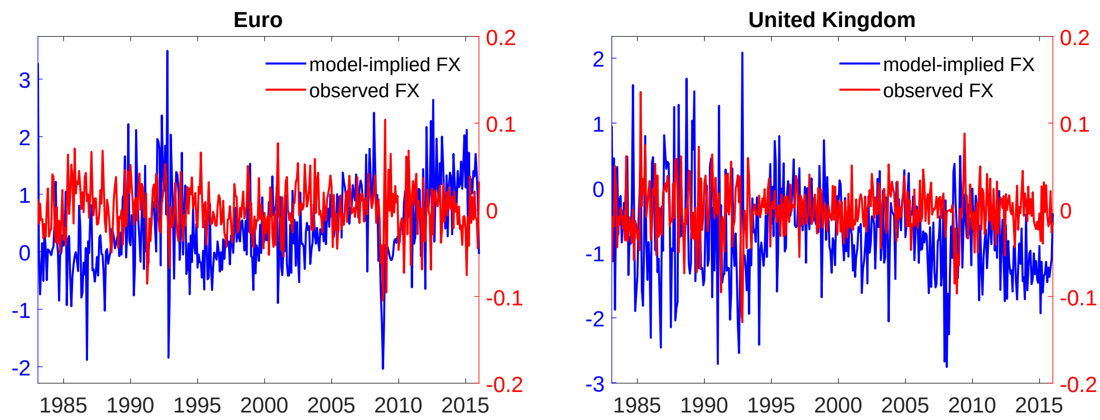
Notes: We plot loadings on depreciation rates that are used by the WFX model to establish U.S. and foreign bond yields for multiple horizons (0-60 month, x-axis). Different lines represent loadings for bonds of different countries. None of the lines are signifciantly different from zero. We do not report confidence intervals to avoid clutter.

Figure 3
Yield factor loadings for the WFX model



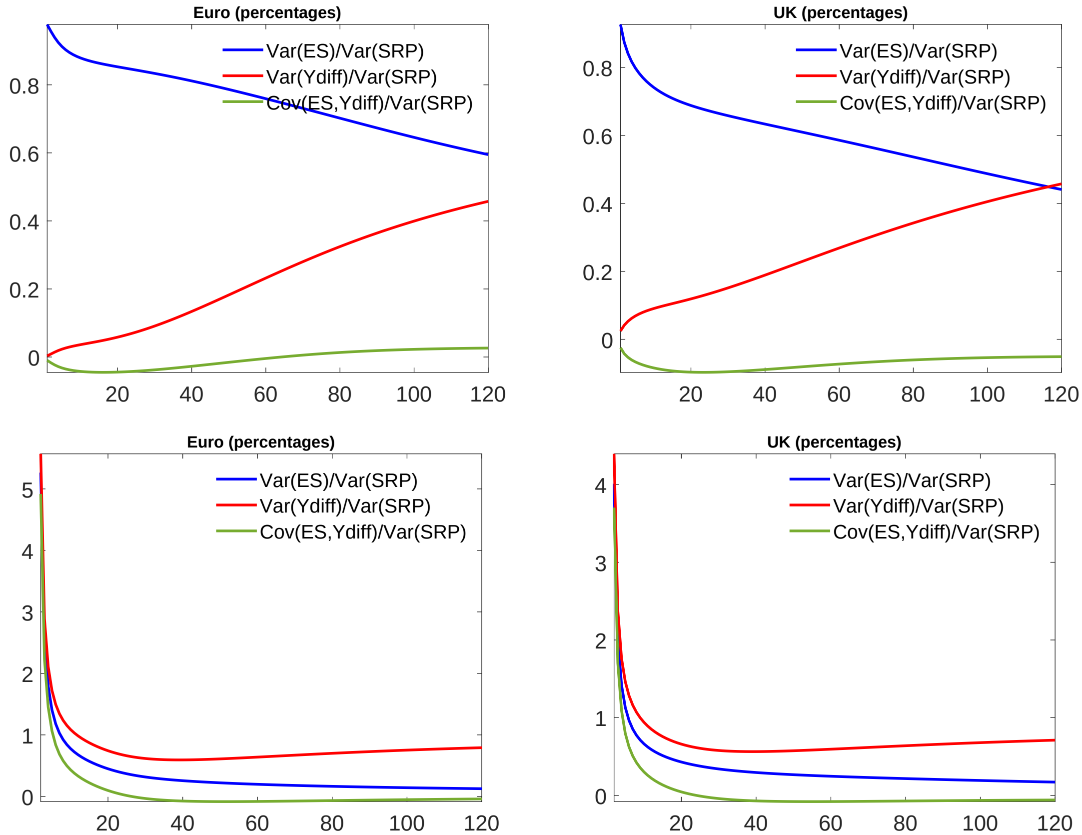
Notes: We plot yield factor loadings that are used by the WFX model to establish U.S. and foreign bond yields for multiple horizons (0-60 month, x-axis). Different lines represent loadings for bonds of different countries.

Figure 4
NFX-implied and observed FX rates



Notes: We plot the depreciation rates $\Delta s_{t+1} = \widehat{m}_{t,t+1}(Q, \widehat{Q}) - m_{t,t+1}(Q, \widehat{Q})$ *implied by the NFX model (blue, vertical axis left) against the observed depreciation rates (red, vertical axis right).*

Figure 5
News-based decomposition of currency risk premiums



Notes: We plot percentage contribution to news about currency risk premiums, according to $1 = var(N^n_{s,t})/var(N^n_{srp,t}) + var(N^n_{\Delta y,t})/var(N^n_{srp,t}) - 2cov(N^n_{s,t}, N^n_{\Delta y,t})/var(N^n_{srp,t})$ *across the different horizons* n*. The first row uses the WFX model to compute the decomposition. The second row uses the NFX model combined with the UIP regression (to establish srp).*

Table 1: Principal components

PC's	6 PCs (US + €)	6 PCs (US + £)	9 PCs (all countries)
1	91.5142	96.1800	93.4792
2	99.3414	99.3000	97.9044
3	99.7994	99.8800	99.3445
4	99.9820	99.9800	99.8226
5	99.9924	99.9960	99.9304
6	99.9995	99.9995	99.9807
7	-	-	99.9923
8	-	-	99.9969
9	-	-	99.9995

We report per cent of variation in international yield curves explained by principal components for various scenarios of data used.

Table 2: R^2 from spanning regressions, %

Panel A. Regression of currency returns on bond and equity returns

FX	Type of R^2	Bond returns		Bond and equity returns	
		\$ returns	€ or £ returns	\$ returns	€ or £ returns
€	R^2	15.65	12.80	28.04	23.08
	R^2_{adj}	14.53	11.65	26.70	21.86
£	R^2	9.59	14.42	26.15	14.42
	R^2_{adj}	8.39	13.28	24.77	13.05

Panel B. Regression of depreciation rate on principal components of yields

FX	Type of R^2	6 PCs (US + country)	6 PCs (all countries)	9 PCs (all countries)
€	R^2	2.90	5.01	6.38
	R^2_{adj}	1.40	3.54	4.20
£	R^2	0.98	2.78	2.92
	R^2_{adj}	-0.55	1.28	0.65

We report R^2, regular and adjusted, expressed in percent for spanning regressions. In panel A we regress annual currency returns of a given country (obtained by investing in a foreign one-period bond) on annual bond returns of maturities n = 2, 3, 4, and 5 years expressed in the same units (USD or foreign). We also combine bond returns with MSCI equity index returns in the last two columns. In panel B we regress monthly depreciation rate of a given country vis-a-vis the USD on principal components constructed from yields on US bonds, bonds of that country, and, in the last column, bonds of the third country as well.

Table 3: Pricing errors across countries

	NFX			WFX		
	US	Euro	UK	US	Euro	UK
y_t^{12}	0.11	0.09	0.10	0.12	0.09	0.10
y_t^{24}	0.08	0.07	0.08	0.09	0.07	0.09
y_t^{36}	0.05	0.05	0.06	0.06	0.05	0.06
y_t^{48}	0.02	0.02	0.03	0.03	0.02	0.03

Posterior mean estimates of the pricing errors in annualized percentage points, $100 \times \sqrt{\operatorname{diag}\left(\Sigma_y \Sigma_y' \times 12\right)}$ *,for the US, Euro, and UK for both the NFX and WFX model. These are reported for yields of different maturity that are not part of the state* x_t*.*

Table 4: Properties of Currency Prices and Interest Rates

Panel A: Summary statistics									
	mean	NFX	WFX	st.dev.	NFX	WFX	autocorr	NFX	WFX
	Δs_t								
Euro	0.909	478.258	0.909	9.99	250.73	9.99	0.150	0.346	0.150
UK	-0.264	-703.672	-0.264	10.18	249.25	10.18	0.055	0.181	0.055
	Short rates								
US	3.725	3.725	3.725	0.79	0.79	0.79	0.985	0.985	0.985
Euro	3.673	3.475	3.459	0.69	0.69	0.70	0.988	0.992	0.992
UK	5.942	5.942	5.942	1.16	1.16	1.16	0.990	0.990	0.990
	$y_t^1 - \widehat{y}_t^1$								
Euro	0.364	0.251	0.266	0.56	0.55	0.55	0.974	0.974	0.974
UK	-2.217	-2.217	-2.217	0.61	0.61	0.61	0.969	0.969	0.969

Panel B: UIP regressions						
$\Delta s_{t+1} = a + b(y_t^1 - \widehat{y}_t^1) + \varepsilon_{t+1}$						
	$\hat{a}$	NFX	WFX	$\hat{b}$	NFX	WFX
Euro	0.0014	0.4123	0.0009	-1.2169	-103.9271	-0.9458
	(0.0017)	(0.0017)	(0.0501)	(1.1654)	(27.1927)	(1.1422)
UK	-0.0016	-0.6529	-0.0016	-0.8387	-33.9749	-0.8387
	(0.0019)	(0.0533)	(0.0019)	(1.1580)	(25.6371)	(1.1580)

We replicate Table I of B/F/T. We report the sample mean, sample standard deviation, and sample autocorrelation.

Table 5: Sample moments of data versus model-implied yields

	mean			st.dev.			autocorr		
	data	NFX	WFX	data	NFX	WFX	data	NFX	WFX
US									
y_t^1	3.725	3.725	3.725	0.791	0.791	0.791	0.985	0.985	0.985
y_t^{12}	4.354	4.160	4.159	0.877	0.848	0.820	0.991	0.989	0.989
y_t^{24}	4.652	4.542	4.534	0.885	0.867	0.839	0.990	0.989	0.989
y_t^{36}	4.907	4.852	4.842	0.874	0.865	0.846	0.989	0.989	0.989
y_t^{48}	5.130	5.109	5.103	0.856	0.853	0.844	0.989	0.989	0.989
y_t^{60}	5.326	5.326	5.326	0.837	0.837	0.837	0.988	0.988	0.988
Euro									
$\widehat{y}_t^1$	3.673	3.476	3.459	0.690	0.693	0.698	0.992	0.992	0.992
$\widehat{y}_t^{12}$	3.817	3.772	3.787	0.736	0.716	0.720	0.991	0.992	0.992
$\widehat{y}_t^{24}$	4.022	4.030	4.047	0.734	0.727	0.727	0.990	0.992	0.992
$\widehat{y}_t^{36}$	4.228	4.243	4.253	0.727	0.725	0.724	0.990	0.991	0.991
$\widehat{y}_t^{48}$	4.413	4.422	4.425	0.716	0.716	0.715	0.989	0.990	0.990
$\widehat{y}_t^{60}$	4.573	4.573	4.573	0.702	0.702	0.702	0.989	0.989	0.989
UK									
$\widehat{y}_t^1$	5.942	5.942	5.942	1.163	1.163	1.163	0.990	0.990	0.990
$\widehat{y}_t^{12}$	5.785	5.841	5.887	1.070	1.087	1.065	0.990	0.991	0.991
$\widehat{y}_t^{24}$	5.895	5.913	5.934	1.028	1.032	1.014	0.990	0.991	0.991
$\widehat{y}_t^{36}$	6.010	6.022	6.022	0.992	0.994	0.984	0.990	0.990	0.990
$\widehat{y}_t^{48}$	6.116	6.123	6.118	0.964	0.966	0.961	0.990	0.990	0.990
$\widehat{y}_t^{60}$	6.208	6.208	6.208	0.942	0.942	0.942	0.989	0.989	0.989

Reduced-form moments vs model-implied moments from the main model.

Appendix A Bond prices

Appendix A.1 U.S. bonds

The price of a one month bond is

$$Q_t^1 \;=\; E_t\left[\exp\left(m_{t,t+1}\right)\right] \;=\; \exp\left(\bar{a}_1 + \bar{b}_{1,x}^{\top} x_t\right)$$

where $\bar{a}_1 \;=\; -\delta_{i,0}$ and $\bar{b}_{1,x} \;=\; -\delta_{i,x}$. The U.S. short rate is

$$i_t \;=\; \delta_{i,0} + \delta_{i,x}^{\top} x_t$$

The price of an n-period U.S. bond is

$$\begin{aligned}
Q_t^n &= E_t\left[\exp\left(m_{t,t+1}\right) Q_{t+1}^{n-1}\right] \;=\; E_t\left[\exp\left(-\delta_{i,0} - \delta_{i,x}^{\top} x_t - \frac{1}{2}\lambda_t^{\top}\lambda_t - \lambda_t^{\top}\varepsilon_{t+1} + \bar{a}_{n-1} + \bar{b}_{n-1,x}^{\top} x_{t+1}\right)\right] \\
&= \exp\left(\bar{a}_{n-1} - \delta_{i,0} - \delta_{i,x}^{\top} x_t - \frac{1}{2}\lambda_t^{\top}\lambda_t\right) E_t\left[\exp\left(-\lambda_t^{\top}\varepsilon_{t+1} + \bar{b}_{n-1,x}^{\top}\left[\mu_x + \Phi_x x_t + \Sigma_x \varepsilon_{t+1}\right]\right)\right] \\
&= \exp\left(\bar{a}_{n-1} - \delta_{i,0} - \delta_{i,x}^{\top} x_t + \bar{b}_{n-1,x}^{\top}\left[\mu_x + \Phi_x x_t\right] - \lambda_t^{\top}\Sigma_x^{\top}\bar{b}_{n-1,x} + \frac{1}{2}\bar{b}_{n-1,x}^{\top}\Sigma_x\Sigma_x^{\top}\bar{b}_{n-1,x}\right) \\
&= \exp\left(\bar{a}_n + \bar{b}_{n,x}^{\top} x_t\right)
\end{aligned}$$

where the loadings are

$$\begin{aligned}
\bar{a}_n &= \bar{a}_{n-1} - \delta_{i,0} + \bar{b}_{n-1,x}^{\top}\left(\mu_x - \lambda_\mu\right) + \frac{1}{2}\bar{b}_{n-1,x}^{\top}\Sigma_x\Sigma_x^{\top}\bar{b}_{n-1,x} \\
\bar{b}_{n,x} &= \left(\Phi_x - \lambda_\phi\right)^{\top}\bar{b}_{n-1,x} - \delta_{i,x}
\end{aligned}$$

We can write this in terms of the U.S. risk neutral parameters

$$\begin{aligned}
\bar{a}_n &= \bar{a}_{n-1} - \delta_{i,0} + \bar{b}_{n-1,x}^{\top}\mu_x^{*} + \frac{1}{2}\bar{b}_{n-1,x}^{\top}\Sigma_x\Sigma_x^{\top}\bar{b}_{n-1,x} \\
\bar{b}_{n,x} &= \Phi_x^{*,\top}\bar{b}_{n-1,x} - \delta_{i,x}
\end{aligned}$$

U.S. yields are $y_t = a_n + b_{n,x}^{\top} x_t$ with $a_n = -n^{-1}\bar{a}_n$ and $b_{n,x} = -n^{-1}\bar{b}_{n,x}$.

Appendix A.2 Foreign bonds in the NFX model

The price of a one month bond is

$$\widehat{Q}_t^1 \;=\; E_t\left[\exp\left(\widehat{m}_{t,t+1}\right)\right] \;=\; \exp\left(\widehat{\bar{a}}_1^N + \widehat{\bar{b}}_{1,x}^{N,\top} x_t\right)$$

where $\widehat{\bar{a}}_1^N \;=\; -\widehat{\delta}_{i,0}$ and $\widehat{\bar{b}}_{1,x}^N \;=\; -\widehat{\delta}_{i,x}$. The foreign short rate is

$$\widehat{i}_t \;=\; \widehat{\delta}_{i,0} + \widehat{\delta}_{i,x}^{\top} x_t$$

Using the same calculations as above, the price of an n-period foreign bond in the NFX model is

$$\begin{aligned}
\widehat{Q}_t^n &= E_t\left[\exp\left(\widehat{m}_{t,t+1}\right)\widehat{Q}_{t+1}^{n-1}\right] \;=\; E_t\left[\exp\left(-\widehat{\delta}_{i,0} - \widehat{\delta}_{i,x}^{\top} x_t - \frac{1}{2}\widehat{\lambda}_t^{\top}\widehat{\lambda}_t - \widehat{\lambda}_t^{\top}\varepsilon_{t+1} + \widehat{\bar{a}}_{n-1}^N + \widehat{\bar{b}}_{n-1,x}^{N,\top} x_{t+1}\right)\right] \\
&= \exp\left(\widehat{\bar{a}}_{n-1}^N - \widehat{\delta}_{i,0} - \widehat{\delta}_{i,x}^{\top} x_t + \widehat{\bar{b}}_{n-1,x}^{N,\top}\left[\mu_x + \Phi_x x_t\right] - \widehat{\lambda}_t^{\top}\Sigma_x^{\top}\widehat{\bar{b}}_{n-1,x}^N + \frac{1}{2}\widehat{\bar{b}}_{n-1,x}^{N,\top}\Sigma_x\Sigma_x^{\top}\widehat{\bar{b}}_{n-1,x}^N\right) \\
&= \exp\left(\widehat{\bar{a}}_n^N + \widehat{\bar{b}}_{n,x}^{N,\top} x_t\right)
\end{aligned}$$

where the loadings are

$$
\begin{aligned}
\widehat{a}_n^N &= \widehat{a}_{n-1}^N - \widehat{\delta}_{i,0} + \widehat{b}_{n-1,x}^{N,\top}\left(\mu_x - \widehat{\lambda}_\mu\right) + \frac{1}{2}\widehat{b}_{n-1,x}^{N,\top}\Sigma_x\Sigma_x^\top\widehat{b}_{n-1,x}^N \\
\widehat{b}_{n,x}^N &= \left(\Phi_x - \widehat{\lambda}_\phi\right)^\top \widehat{b}_{n-1,x}^N - \widehat{\delta}_{i,x}
\end{aligned}
$$

We can write this in terms of the foreign risk neutral parameters

$$
\begin{aligned}
\widehat{a}_n^N &= \widehat{a}_{n-1}^N - \widehat{\delta}_{i,0} + \widehat{b}_{n-1,x}^{N,\top}\widehat{\mu}_x^* + \frac{1}{2}\widehat{b}_{n-1,x}^{N,\top}\Sigma_x\Sigma_x^\top\widehat{b}_{n-1,x}^N \\
\widehat{b}_{n,x}^N &= \widehat{\Phi}_x^{*,\top}\widehat{b}_{n-1,x}^N - \widehat{\delta}_{i,x}
\end{aligned}
$$

Foreign yields are $\widehat{y}_t = \widehat{a}_n^N + \widehat{b}_{n,x}^{N,\top}x_t$ with $\widehat{a}_n^N = -n^{-1}\widehat{a}_n^N$ and $\widehat{b}_{n,x}^N = -n^{-1}\widehat{b}_{n,x}^N$.

Appendix A.3 Foreign bonds in the WFX model

The price of a one month bond is

$$
\widehat{Q}_t^1 = E_t\left[\exp\left(m_{t,t+1} + \Delta s_{t+1}\right)\right] = \exp\left(\widehat{a}_1^W + \widehat{b}_{1,x}^{W,\top}x_t\right)
$$

where $\widehat{a}_1^W = \delta_{s,0} - \delta_{i,0} + \delta_{s,x}^\top\left(\mu_x - \lambda_\mu\right) + \frac{1}{2}\delta_{s,x}^\top\Sigma_x\Sigma_x^\top\delta_{s,x}$ and $\widehat{b}_{1,x}^W = \left(\Phi_x - \lambda_\phi\right)^\top\delta_{s,x} - \delta_{i,x}$. Using the same calculations as above, the price of an n-period foreign bond in the WFX model is

$$
\begin{aligned}
\widehat{Q}_t^n &= E_t\left[\exp\left(m_{t,t+1} + \Delta s_{t+1}\right)\widehat{Q}_{t+1}^{n-1}\right] \\
&= \exp\left(\widehat{a}_{n-1}^W + \delta_{s,0} - \delta_{i,0} - \delta_{i,x}^\top x_t + \left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right)^\top\left[\mu_x + \Phi_x x_t\right]\right) \\
&\quad \exp\left(-\lambda_t^\top\Sigma_x^\top\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right) + \frac{1}{2}\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right)^\top\Sigma_x\Sigma_x^\top\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right)\right) \\
&= \exp\left(\widehat{a}_n^W + \widehat{b}_{n,x}^{W,\top}x_t\right)
\end{aligned}
$$

where the loadings are

$$
\begin{aligned}
\widehat{a}_n^W &= \widehat{a}_{n-1}^W - \delta_{i,0} + \left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right)^\top\left(\mu_x - \lambda_\mu\right) + \frac{1}{2}\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right)^\top\Sigma_x\Sigma_x^\top\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right) \\
\widehat{b}_{n,x}^W &= \left(\Phi_x - \lambda_\phi\right)^\top\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right) - \delta_{i,x}
\end{aligned}
$$

We can write this in terms of the U.S. risk neutral parameters

$$
\begin{aligned}
\widehat{a}_n^W &= \widehat{a}_{n-1}^W - \delta_{i,0} + \left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right)^\top\mu_x^* + \frac{1}{2}\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right)^\top\Sigma_x\Sigma_x^\top\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right) \\
\widehat{b}_{n,x}^W &= \Phi_x^{*,\top}\left(\widehat{b}_{n-1,x}^W + \delta_{s,x}\right) - \delta_{i,x}
\end{aligned}
$$

Foreign yields are $\widehat{y}_t = \widehat{a}_n^W + \widehat{b}_{n,x}^{W,\top}x_t$ with $\widehat{a}_n^W = -n^{-1}\widehat{a}_n^W$ and $\widehat{b}_{n,x}^W = -n^{-1}\widehat{b}_{n,x}^W$.

Appendix B Estimation

Appendix B.1 Parameterization and identification

In the main text, we report estimates of the market price of risk parameters λ_0 and λ_x that enter the stochastic discount factor. These parameters are defined in terms of the observable state vector x_t as defined in the text. In practice, these parameters λ_0 and λ_x require identifying restrictions that are not easy to impose directly. The term structure literature solves the problem of imposing the necessary identifying restrictions by parameterizing the model in terms of the identifiable risk neutral parameters under a latent factor rotation. Under the latent factor rotation, the restrictions are easy to impose. We follow this literature and explain how to impose these restrictions in this appendix.

First, we note that the risk neutral parameters μ_x^* and Φ_x^* of the observable state vector x_t are related to the market prices of risk as

$$\begin{aligned}\mu_x^* &= \mu_x - \lambda_0 \\ \Phi_x^* &= \Phi_x - \lambda_x\end{aligned}$$

where μ_x and Φ_x are the drift and autocovariance of x_t under the "real-world" probabilities.

In this appendix, we use a 'tilde' to denote any parameters $\tilde{\theta}$ or state variables $\tilde{x}_t$ under the latent factor rotation. In our implementation, we make one minor change relative to the term structure literature. We define $\tilde{x}_t$ such that the first two elements are the observed depreciation rates while the remaining yield factors are latent, i.e. we have

$$\tilde{x}_t = \begin{pmatrix} \Delta s_t \\ \tilde{g}_t \end{pmatrix}$$

where $\tilde{g}_t$ are latent yield factors. Therefore, when we rotate from $\tilde{x}_t$ to x_t, the first two elements of the state remain the same.

With this definition of $\tilde{x}_t$ in hand, the observed factors x_t are related to the latent factors $\tilde{x}_t$ through the linear transformation

$$x_t = \Gamma_0 + \Gamma_1 \tilde{x}_t$$

where the vector Γ_0 and matrix Γ_1 are described below in Appendix B.3.

The risk neutral dynamics under the latent factor rotation are

$$\begin{aligned}\Delta s_t &= \tilde{\delta}_{s,0} + \tilde{\delta}_{s,x}^{\top} \tilde{x}_t \\ i_t &= \tilde{\delta}_{i,0} + \tilde{\delta}_{i,x}^{\top} \tilde{x}_t \\ \tilde{x}_t &= \tilde{\mu}_x^* + \tilde{\Phi}_x^* \tilde{x}_{t-1} + \tilde{\Sigma}_x \varepsilon_t^*\end{aligned}$$

Under this rotation, the identifying restrictions require imposing the following restrictions

$$\begin{aligned}\tilde{\delta}_{s,0} &= 0 \\ \tilde{\delta}_{s,x} &= \mathrm{e}_i \qquad i = 1,2 \\ \tilde{\delta}_{i,x} &= \begin{pmatrix} \tilde{\delta}_{i,s} \\ \iota \end{pmatrix} \\ \tilde{\mu}_x^* &= \begin{pmatrix} \tilde{\mu}_s^* \\ \tilde{\mu}_g^* \end{pmatrix} = \begin{pmatrix} \tilde{\mu}_s^* \\ 0 \end{pmatrix} \\ \tilde{\Phi}_x^* &= \begin{pmatrix} \tilde{\Phi}_s^* & \tilde{\Phi}_{sg}^* \\ \tilde{\Phi}_{gs}^* & \tilde{\Phi}_g^* \end{pmatrix} = \begin{pmatrix} \tilde{\Phi}_s^* & \tilde{\Phi}_{sg}^* \\ 0 & \tilde{\Phi}_g^* \end{pmatrix}\end{aligned}$$

In addition, the matrix $\tilde{\Phi}_g^*$ is restricted to be a matrix of eigenvalues. In general, the eigenvalues may be distinct and real, complex, or repeated. We follow the standard approach in the term structure literature and assume that this matrix is diagonal with distinct, real eigenvalues ordered from largest to smallest.

For the loadings on the U.S. short rate $\tilde{\delta}_{i,x}$, the first 2 elements associated with depreciation rate are estimable while the remaining loadings are restricted to be a vector of ones ι.

Appendix B.2 Observables

We stack the U.S. and foreign nominal yields of different maturities into vectors $y_t = \left(y_t^1, \ldots, y_t^{60}\right)$ and $\widehat{y}_t = \left(\widehat{y}_t^1, \ldots, \widehat{y}_t^{60}\right)$ as well as their bond loadings, e.g. $A = (a_1, \ldots, a_{60})^\top, B = (b_{1,x}, \ldots, b_{60,x})^\top$ and $\widehat{A} = (\widehat{a}_1, \ldots, \widehat{a}_{60})^\top, \widehat{B} = \left(\widehat{b}_{1,x}, \ldots, \widehat{b}_{60,x}\right)^\top$.

The system of observation equations used in the model are

$$\begin{aligned}
\Delta s_t &= \delta_{s,0} + \delta_{s,x} x_t \\
y_t &= A + B x_t \\
\widehat{y}_t &= \widehat{A} + \widehat{B} x_t
\end{aligned}$$

where, in our application, the vector $\widehat{y}_t$ includes both Euro and U.K. yields. We define the overall vector of observables as

$$Y_t = \begin{pmatrix} \Delta s_t \\ y_t \\ \widehat{y}_t \end{pmatrix}$$

Next, we define two selection matrices W_1 and W_2 that select out of Y_t linear combinations of depreciation rates and yields. Together, the matrices $\left(W_1^\top; W_2^\top\right)^\top$ must be full rank. These linear combinations are

$$\begin{aligned}
Y_t^{(1)} &= W_1 Y_t \\
Y_t^{(2)} &= W_2 Y_t
\end{aligned}$$

The vector $Y_t^{(1)}$ is a linear combination of observables that we assume to be measured without error. The vector $Y_t^{(2)}$ is observed with error.

Appendix B.3 Rotating the state vector to observables

In our implementation, we choose W_1 so that the state vector is $x_t = Y_t^1 = W_1 Y_t$ with x_t defined as in the text

$$x_t = \begin{pmatrix} \Delta s_t^{€} \\ \Delta s_t^{£} \\ y_t^1 \\ y_t^1 - \widehat{y}_t^{€,1} \\ y_t^1 - \widehat{y}_t^{£,1} \\ y_t^{60,1} \\ y_t^{60,1} - \widehat{y}_t^{€,60,1} \\ y_t^{60,1} - \widehat{y}_t^{£,60,1} \end{pmatrix}$$

In order for x_t to have exactly this definition, we follow the term structure literature and assume that this linear combination of the observables Y_t is observed without measurement error.

The matrix W_2 is a selection matrix full of zeros and ones that selects unique linear combinations out of Y_t that are not used in x_t. Specifically, W_2 is defined such that the second set of linear combinations $Y_t^{(2)} = W_2 Y_t$ includes the 12, 24, 36, and 48 month U.S. yields as well as the 12, 24, 36, and 48 month foreign yields (both Euro and U.K.).

To implement this rotation in practice, we note that the observables Y_t are related to the latent state vector $\tilde{x}_t$ as

$$Y_t = \begin{pmatrix} \Delta s_t \\ y_t \\ \widehat{y}_t \end{pmatrix} = \begin{pmatrix} \tilde{\delta}_{s,0} \\ \tilde{A} \\ \tilde{\widehat{A}} \end{pmatrix} + \begin{pmatrix} \tilde{\delta}_{s,x}^{\top} \\ \tilde{B} \\ \tilde{\widehat{B}} \end{pmatrix} \tilde{x}_t = \tilde{C} + \tilde{D}\tilde{x}_t$$

where the vector $\tilde{C}$ and matrix $\tilde{D}$ are appropriately defined. In this case, the bond loadings $\tilde{A}, \tilde{\widehat{A}}, \tilde{B}, \tilde{\widehat{B}}$ are calculated under the latent factor rotation. Next, we pre-multiply Y_t above by W_1 and substitute in for the observed state variables

$$\begin{aligned} W_1 Y_t &= W_1 \tilde{C} + W_1 \tilde{D} \tilde{x}_t \\ &= W_1 \tilde{C} + W_1 \tilde{D} \Gamma_1^{-1} (x_t - \Gamma_0) \\ &= W_1 \left(\tilde{C} - \tilde{D} \Gamma_1^{-1} \Gamma_0 \right) + W_1 \tilde{D} \Gamma_1^{-1} x_t. \end{aligned}$$

In order for $x_t = W_1 Y_t$, it implies that the rotation matrices Γ_0 and Γ_1 must satisfy the restrictions

$$\begin{aligned} \Gamma_0 &= W_1 \tilde{C} \\ \Gamma_1 &= W_1 \tilde{D} \end{aligned}$$

Given these matrices, we can map between the parameters of the observable rotation of the state vector x_t and the latent factor rotation $\tilde{x}_t$ through the transformations

$$\begin{aligned} \Phi_x^* &= \Gamma_1 \tilde{\Phi}_x^* \Gamma_1^{-1} \\ \mu_x^* &= (I - \Phi_x^*) \Gamma_0 + \Gamma_1 \tilde{\mu}_x^* \end{aligned}$$

Appendix B.4 Prior distributions

- Let $S_y = \Sigma_y \Sigma_y^{\top}$ with dimension $d_{y_2} \times d_{y_2}$. Note that $Y_t^{(2)}$ has dimension $d_{y_2} \times 1$. We assume S_y has a diffuse inverse Wishart distribution $S_y \sim \text{Inv-W}\left(\underline{\Omega}_y, \underline{\nu}_y\right)$ with degrees of freedom $\underline{\nu}_y = 0$ and scale matrix $\underline{\Omega}_y = 0$.
- The matrix Σ_x is lower triangular. We place inverse Gamma $\sigma_i^2 \sim \text{IG}(\alpha_i, \beta_i)$ on each of the diagonal elements where $i = s, g$. The subscript s stands for depreciation rate while the subscript g stands for yield factor. We set $\alpha_s = 3.3$ and $\beta_s = 0.0015$. We set $\alpha_g = 3.05$ and $\beta_g = 8e^{-8}$.
- We place a prior on the free parameters of the unconditional means $\left(\bar{\mu}_x, \bar{\tilde{\mu}}_x^*\right)$ directly instead of the drifts $(\mu_x, \tilde{\mu}_x^*)$. First, we calculate the unconditional sample mean of the factors $\hat{\bar{\mu}}_x$. Our prior for each element of $\bar{\mu}_x$ is a normal distribution centered at the sample mean. Then, we choose the variance of this distribution to be large enough to cover the support of the data. Our priors are
  - depreciation rates: $\bar{\mu}_s \sim \text{N}\left(\hat{\bar{\mu}}_s, 0.0003\right)$
  - U.S. level: $\bar{\mu}_i \sim \text{N}\left(\hat{\bar{\mu}}_i, 0.000003\right)$
  - Foreign Level differentials: $\bar{\mu}_{\Delta_c i} \sim \text{N}\left(\hat{\bar{\mu}}_{\Delta_c i}, 0.000003\right)$
  - U.S. slope: $\bar{\mu}_{sl} \sim \text{N}\left(\hat{\bar{\mu}}_{sl}, 0.000003\right)$
  - Foreign slope differential: $\bar{\mu}_{\Delta_c sl} \sim \text{N}\left(\hat{\bar{\mu}}_{\Delta_c sl}, 0.000003\right)$

Identifying restrictions on $\bar{\tilde{\mu}}_x^*$ only allow for only 2 free parameters, which are the parameters associated with the depreciation rates (i.e. the first two factors in $\tilde{x}_t$). Our prior for this variable is the same as the unconditional depreciation rate under the real world probability but multiplied by a factor of 100.

- depreciation rates: $\bar{\tilde{\mu}}_s \sim \mathrm{N}\left(\hat{\bar{\mu}}_s, 0.03\right)$

- We parameterize the matrix $\tilde{\Phi}_x^*$ as

$$\tilde{\Phi}_x^* = \begin{pmatrix} \tilde{\Phi}_s^* & \tilde{\Phi}_{sg}^* \\ 0 & \tilde{\Phi}_g^* \end{pmatrix}$$

Our priors on the sub-matrices are as follows:

  - $\tilde{\Phi}_g^*$ is a diagonal matrix of real, ordered eigenvalues. Let $a_1 = -1$ and $b = 1$. We parameterize them as $\tilde{\Phi}_{g,11}^* = a_1 + (b - a_1)U_1$ and $\tilde{\Phi}_{g,jj}^* = a_{j-1} + (b - a_{j-1})U_j$ for $j = 2, \ldots, d_g$. This transformation ensures that they are increasing and contained in the interval $[-1, 1]$. We then place priors on $\tilde{\Phi}_{g,jj}^*$ via $U_j \sim \mathrm{Beta}\,(12, 12)$.
  - We place the same prior on $\tilde{\Phi}_s^*$ and $\tilde{\Phi}_{sg}^*$ as under the risk neutral dynamics but where the covariance matrix is multiplied by a factor of 100.

- We separate $\tilde{\delta}_{i,x}$ into two sub-vectors $\tilde{\delta}_{i,s}$ and $\tilde{\delta}_{i,g}$. We place a prior on the free parameters of the factor loadings $\tilde{\delta}_{i,s}$. Our identifying restriction is that $\tilde{\delta}_{i,g} = \iota$. The parameters of $\tilde{\delta}_{i,s}$ are estimable and we assume that each entry is independent and distributed as $\tilde{\delta}_{i,g} \sim \mathrm{N}\,(0, 0.01)$.

Appendix B.5 Log-likelihood function

The log-likelihood function is

$$\mathcal{L} = \log p\,(Y_1, \ldots, Y_T|\theta) = \sum_{t=1}^{T} \log p\,(x_t|x_{t-1}, \theta) + \sum_{t=1}^{T} \log p\left(Y_t^{(2)}|x_t; \theta\right)$$

where x_0 are assumed to be known. The density $p\,(x_t|x_{t-1}; \theta)$ is determined by the VAR dynamics of the factors x_t while the second term comes from the linear combination of yields observed with error

$$Y_t^{(2)} = C^{(2)} + D^{(2)} x_t + \Sigma_y \eta_t, \qquad \eta_t \sim \mathrm{N}\,(0, \mathrm{I}),$$

where $C^{(2)} = W_2 C$ and $D^{(2)} = W_2 D$ and

$$\begin{aligned} C &= \tilde{C} - \tilde{C}\Gamma_1^{-1}\Gamma_0, \\ D &= \tilde{D}\Gamma_1^{-1}. \end{aligned}$$

This likelihood function assumes that there are no missing values in either $Y_t^{(1)}$ or $Y_t^{(2)}$. In practice, this is not the case. We impute these missing values during the MCMC algorithm using the Kalman filter.

Appendix B.6 Estimation

Let θ denote all the parameters of the model and define $f_{1:T} = (f_1, \ldots, f_T)$ and $Y_{1:T} = (Y_1, \ldots, Y_T)$. In practice, some data points are missing which implies that some of the factors f_t are missing. We use $Y_{1:T}^o$ and $Y_{1:T}^m$ to denote the observed and missing data, respectively. The joint posterior distribution over the parameters and missing data is given by

$$p\,(\theta, Y_{1:T}^m|Y_{1:T}^o) \propto p\,(Y_{1:T}^o|\theta)\,p\,(\theta),$$

where $p\,(Y_{1:T}^o|\theta)$ is the likelihood and $p\,(\theta)$ is the prior distribution. We use Markov-chain Monte Carlo to draw from the posterior.

Appendix B.6.1 MCMC algorithm

We provide a brief description of the MCMC algorithm. Let $S_y = \Sigma_y \Sigma_y'$ and $S_x = \Sigma_x \Sigma_x'$ denote the covariance matrices. We use a Gibbs sampler that iterates between drawing from each of the full conditional distributions.

- Place the model in linear, Gaussian state space form as described in Appendix B.6.2. Draw the missing data and unconditional means $(Y^m_{1:T}, \bar{\mu}_x, \bar{\mu}^*_x)$ from their full conditional distribution using the Kalman filter and simulation smoothing algorithm. Given the full data $Y^{o,m}_t = (Y^o_t, Y^m_t)$, we can recalculate the factors $x_t = W_1 Y^{o,m}_t$.
- Let $\bar{x}_t = x_t - \bar{\mu}_x$ denote the demeaned factors. We draw the free elements of Φ_x from their full conditional distribution using standard results for Bayesian multiple regression. We write the VAR as a regression model
$$\bar{x}_t = X_t \phi_x + \Sigma_x \varepsilon_t$$
where $\phi_x = \text{vec}(\Phi_x)$ and the regressors X_t contain lagged values of $\bar{x}_{t-1}$. Draws from this model are standard.
- Draw the free elements of S_x from their full conditional using a random-walk Metropolis algorithm. In this step, we avoid conditioning on the parameters S_y, Φ_x by analytically integrating these parameters out of the likelihood.
- Draw the eigenvalues Λ^*_x from their full conditional using random-walk Metropolis. To avoid conditioning on S_y, Φ_x, we draw from the marginal distribution that analytically integrates these values out of the likelihood.
- Draw the elements of $\tilde{\delta}_{s,x}$ from their full conditional using random-walk Metropolis. To avoid conditioning on S_y, Φ_x, we draw from the marginal distribution that analytically integrates these values out of the likelihood.
- The full conditional posterior of S_y is an inverse Wishart distribution $S_y \sim \text{Inv-Wish}\left(\bar{\nu}, \bar{\Omega}\right)$ where $\bar{\nu} = \underline{\nu} + T$ and $\bar{\Omega} = \underline{\Omega} + \sum_{t=1}^{T} \eta_t \eta_t^\top$.

Appendix B.6.2 State space form

In our data set, some of the yields contain missing values. We impute them using the Kalman filter. Given that $x_t = Y^{(1)}_t$, we can write the model in VAR form as

$$\begin{pmatrix} Y^{(1)}_t \\ Y^{(2)}_t \end{pmatrix} = \begin{pmatrix} \mu_x \\ A^{(2)} + B^{(2)}\mu_x \end{pmatrix} + \begin{pmatrix} \Phi_x & 0 \\ B^{(2)}\Phi_x & 0 \end{pmatrix} \begin{pmatrix} Y^{(1)}_{t-1} \\ Y^{(2)}_{t-1} \end{pmatrix} + \begin{pmatrix} \Sigma_x & 0 \\ B^{(2)}\Sigma_x & \Sigma_y \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}$$

Next we translate this system back into Y_t using the fact that

$$Y_t = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} Y^{(1)}_t \\ Y^{(2)}_t \end{pmatrix}$$

to get

$$\begin{aligned} Y_t = & \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \mu_x \\ A^{(2)} + B^{(2)}\mu_x \end{pmatrix} + \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \Phi_x & 0 \\ B^{(2)}\Phi_x & 0 \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} Y_{t-1} \\ & + \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \Sigma_x & 0 \\ B^{(2)}\Sigma_x & \Sigma_y \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \end{aligned}$$

This structure implies that Y_t is a reduced-rank VAR of the form

$$Y_t = \mu_Y + \Phi_Y Y_{t-1} + \Sigma_Y \varepsilon_{Y,t} \qquad \varepsilon_{Y,t} \sim \text{N}(0, I)$$

where

$$\mu_Y = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \mu_x \\ A^{(2)} + B^{(2)}\mu_x \end{pmatrix} \qquad \Phi_Y = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \Phi_x & 0 \\ B^{(2)}\Phi_x & 0 \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}$$

$$\Sigma_Y = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \Sigma_x & 0 \\ B^{(2)}\Sigma_x & \Sigma_y \end{pmatrix} \qquad \varepsilon_{Y,t} = \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix}$$

We place this model in the following linear, Gaussian state space form

$$Y_t = Z\alpha_t + d + u_t \qquad u_t \sim \mathrm{N}(0, H), \tag{B.1}$$

$$\alpha_{t+1} = T\alpha_t + c + Rv_t \qquad v_t \sim \mathrm{N}(0, Q). \tag{B.2}$$

where the initial condition is $\alpha_1 \sim \mathrm{N}\left(a_{1|0}, P_{1|0}\right)$.

Let $\bar{\mu} = \left(\bar{\mu}_{x,u}^{\top}\ \bar{\mu}_{x,u}^{*,\top}\ \delta_{\hat{c},0}\right)^{\top}$ denote the vector of unrestricted unconditional means that enter $\bar{\mu}_x$ and $\bar{\mu}_x^*$ plus the intercept $\delta_{\hat{c},0}$. The vector of intercepts μ_Y can be written as a linear function of the unconditional means

$$\mu_Y = S_{\mu,0} + S_{\mu,1}\bar{\mu}$$

We draw unconditional means jointly with the missing data by including them in the state vector. We define the system matrices from (B.1)-(B.2) as

$$d = 0 \qquad Z = \begin{pmatrix} \mathrm{I} & 0 \end{pmatrix} \qquad H = 0 \qquad Q = \Sigma_Y \Sigma_Y^{\top}$$

$$\alpha_t = \begin{pmatrix} Y_t \\ \bar{\mu} \end{pmatrix} \qquad T = \begin{pmatrix} \Phi_Y & S_{\mu,1} \\ 0 & \mathrm{I} \end{pmatrix} \qquad c = \begin{pmatrix} S_{\mu,0} \\ 0 \end{pmatrix} \qquad R = \begin{pmatrix} \mathrm{I} \\ 0 \end{pmatrix}$$

$$a_{1|0} = \begin{pmatrix} S_{\mu,1}\bar{m}_\mu \\ \bar{m}_\mu \end{pmatrix} \qquad P_{1|0} = \begin{pmatrix} \Sigma_Y\Sigma_Y^{\top} + S_{\mu,1}V_\mu S_{\mu,1}^{\top} & S_{\mu,1}V_\mu \\ V_\mu S_{\mu,1}^{\top} & V_\mu \end{pmatrix}$$

where the prior on the unconditional means is $\bar{\mu} \sim \mathrm{N}(\bar{m}_\mu, V_\mu)$. We use the Kalman filter and simulation smoothing algorithm to draw the missing values and the unconditional means jointly.

Appendix C Estimated parameters in the affine models

Table Appendix C.1: Parameter estimates, state dynamics; NFX model.

x_t^N		x_{1t}	x_{2t}	x_{3t}	x_{4t}	x_{5t}	x_{6t}
	$\bar{\mu}_x \times 1200$	Φ_x					
x_{1t}	3.814	0.976	-0.005	0.045	0.075	-0.051	0.064
	(1.000)	(1.000)	(0.563)	(0.865)	(0.959)	(0.796)	(0.870)
x_{2t}	0.389	-0.025	1.052	-0.029	0.024	0.137	-0.062
	(0.761)	(0.986)	(1.000)	(0.829)	(0.771)	(0.998)	(0.927)
x_{3t}	-2.262	-0.021	-0.029	1.037	0.054	-0.022	0.110
	(0.999)	(0.923)	(0.805)	(1.000)	(0.910)	(0.638)	(0.978)
x_{4t}	1.603	0.008	-0.052	0.009	0.955	-0.117	0.015
	(1.000)	(0.750)	(0.971)	(0.609)	(1.000)	(0.991)	(0.635)
x_{5t}	0.477	0.020	-0.118	0.062	0.060	0.694	0.099
	(0.980)	(0.918)	(1.000)	(0.952)	(0.937)	(1.000)	(0.972)
x_{6t}	1.403	0.002	0.081	-0.146	-0.093	0.089	0.752
	(1.000)	(0.550)	(0.997)	(1.000)	(0.994)	(0.951)	(1.000)
	$\delta_{i,x}$	$\Sigma_x \times \sqrt{12} \times 100$					
x_{1t}	0	0.161	0	0	0	0	0
	(—)	(1.000)	(—)	(—)	(—)	(—)	(—)
x_{2t}	0	0.098	0.076	0	0	0	0
	(—)	(1.000)	(1.000)	(—)	(—)	(—)	(—)
x_{3t}	1	0.029	0.083	0.125	0	0	0
	(—)	(1.000)	(1.000)	(1.000)	(—)	(—)	(—)
x_{4t}	0	-0.034	-0.064	-0.015	0.100	0	0
	(—)	(1.000)	(1.000)	(0.999)	(1.000)	(—)	(—)
x_{5t}	0	-0.057	-0.098	0.009	0.066	0.070	0
	(—)	(1.000)	(1.000)	(0.958)	(1.000)	(1.000)	(—)
x_{6t}	0	-0.029	-0.067	-0.084	0.045	0.012	0.068
	(—)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)

Posterior mean and probability that absolute value of the parameter is greater than zero (in parenthesis) of $\bar{\mu}_x, \Phi_x, \Sigma_x$ from the NFX model. The state variables are: $x_{1t} = y_t^1$, $x_{2t} = y_t^1 - \widehat{y}_t^{€,1}$, $x_{3t} = y_t^1 - \widehat{y}_t^{£,1}$, $x_{4t} = y_t^{60} - y_t^1$, $x_{5t} = \left(y_t^{60} - y_t^1\right) - \left(\widehat{y}_t^{€,60} - \widehat{y}_t^{€,1}\right)$, $x_{6t} = \left(y_t^{60} - y_t^1\right) - \left(\widehat{y}_t^{£,60} - \widehat{y}_t^{£,1}\right)$.

Table Appendix C.2: Paramter estimates, risk premiums; NFX model.

x_t^N		x_{1t}	x_{2t}	x_{3t}	x_{4t}	x_{5t}	x_{6t}
	$\lambda_0 \times 1200$		λ_x				
x_{1t}	0.154	-0.038	-0.005	0.032	-0.046	-0.012	0.040
	(0.990)	(0.997)	(0.559)	(0.787)	(0.861)	(0.573)	(0.764)
x_{2t}	-0.279	-0.018	0.055	-0.026	0.052	0.053	-0.063
	(1.000)	(0.935)	(0.982)	(0.808)	(0.945)	(0.870)	(0.931)
x_{3t}	0.037	-0.023	-0.026	0.039	0.043	-0.017	0.016
	(0.681)	(0.941)	(0.779)	(0.838)	(0.860)	(0.610)	(0.614)
x_{4t}	-0.057	0.024	-0.053	0.024	0.059	-0.158	0.043
	(0.855)	(0.978)	(0.973)	(0.777)	(0.961)	(1.000)	(0.834)
x_{5t}	0.083	0.017	-0.114	0.063	0.058	-0.250	0.097
	(0.902)	(0.889)	(1.000)	(0.955)	(0.929)	(1.000)	(0.968)
x_{6t}	0.089	0.003	0.081	-0.144	-0.088	0.087	-0.165
	(0.924)	(0.593)	(0.998)	(1.000)	(0.992)	(0.946)	(1.000)
	$\widehat{\lambda}_0^{€} \times 1200$		$\widehat{\lambda}_x^{€}$				
x_{1t}	-0.203	-0.029	0.012	0.041	0.005	0.005	0.049
	(0.993)	(0.978)	(0.638)	(0.844)	(0.553)	(0.534)	(0.807)
x_{2t}	-0.288	-0.022	0.064	-0.027	0.035	0.109	-0.058
	(1.000)	(0.972)	(0.993)	(0.816)	(0.858)	(0.991)	(0.915)
x_{3t}	0.017	-0.023	-0.018	0.042	0.043	-0.002	0.063
	(0.568)	(0.940)	(0.710)	(0.855)	(0.858)	(0.518)	(0.875)
x_{4t}	0.060	0.017	-0.057	0.016	0.016	-0.152	0.036
	(0.866)	(0.921)	(0.982)	(0.696)	(0.677)	(1.000)	(0.792)
x_{5t}	0.054	0.019	-0.117	0.063	0.059	-0.278	0.097
	(0.804)	(0.907)	(1.000)	(0.953)	(0.934)	(1.000)	(0.969)
x_{6t}	0.087	0.003	0.080	-0.145	-0.088	0.085	-0.201
	(0.926)	(0.585)	(0.997)	(1.000)	(0.992)	(0.943)	(1.000)
	$\widehat{\lambda}_0^{£} \times 1200$		$\widehat{\lambda}_x^{£}$				
x_{1t}	0.130	-0.044	0.013	0.033	-0.087	0.022	0.028
	(0.942)	(0.999)	(0.646)	(0.793)	(0.977)	(0.638)	(0.690)
x_{2t}	-0.434	-0.016	0.062	-0.025	0.058	0.038	-0.060
	(1.000)	(0.914)	(0.990)	(0.799)	(0.962)	(0.792)	(0.921)
x_{3t}	-0.091	-0.025	-0.015	0.038	0.033	0.004	-0.010
	(0.865)	(0.952)	(0.667)	(0.835)	(0.790)	(0.526)	(0.570)
x_{4t}	-0.061	0.029	-0.057	0.028	0.089	-0.173	0.060
	(0.850)	(0.992)	(0.982)	(0.810)	(0.996)	(1.000)	(0.907)
x_{5t}	0.112	0.017	-0.113	0.064	0.057	-0.233	0.097
	(0.955)	(0.877)	(1.000)	(0.956)	(0.924)	(1.000)	(0.968)
x_{6t}	0.088	0.004	0.081	-0.143	-0.084	0.082	-0.141
	(0.918)	(0.620)	(0.997)	(1.000)	(0.989)	(0.936)	(0.997)

Posterior mean and probability that absolute value of the parameter is greater than zero (in parenthesis). The state variables are: $x_{1t} = y_t^1$, $x_{2t} = y_t^1 - \widehat{y}_t^{€,1}$, $x_{3t} = y_t^1 - \widehat{y}_t^{£,1}$, $x_{4t} = y_t^{60} - y_t^1$, $x_{5t} = \left(y_t^{60} - y_t^1\right) - \left(\widehat{y}_t^{€,60} - \widehat{y}_t^{€,1}\right)$, $x_{6t} = \left(y_t^{60} - y_t^1\right) - \left(\widehat{y}_t^{£,60} - \widehat{y}_t^{£,1}\right)$.

Table Appendix C.3: Parameter estimates; WFX model.

x_t^W		$\Delta s_t^{€}$	$\Delta s_t^{£}$	x_{1t}	x_{2t}	x_{3t}	x_{4t}	x_{5t}	x_{6t}
	$\bar{\mu}_x \times 1200$		Φ_x						
$\Delta s_t^{€}$	0.974	0.019	0.204	-0.696	0.725	-3.710	-0.230	0.274	-2.029
	(0.639)	(0.636)	(1.000)	(0.783)	(0.664)	(0.972)	(0.543)	(0.544)	(0.775)
$\Delta s_t^{£}$	-0.019	0.061	0.010	-0.869	0.510	-0.678	2.211	-4.172	1.589
	(0.504)	(0.856)	(0.575)	(0.826)	(0.614)	(0.636)	(0.841)	(0.919)	(0.714)
x_{1t}	3.802	-4.61e-04	4.57e-04	0.976	1.39e-04	0.038	0.068	-0.036	0.054
	(1.000)	(0.688)	(0.691)	(1.000)	(0.503)	(0.825)	(0.949)	(0.724)	(0.833)
x_{2t}	0.395	-6.73e-05	6.86e-05	-0.026	1.055	-0.033	0.022	0.142	-0.067
	(0.769)	(0.536)	(0.540)	(0.988)	(1.000)	(0.852)	(0.753)	(0.999)	(0.940)
x_{3t}	-2.248	-0.001	5.70e-04	-0.022	-0.026	1.030	0.052	-0.018	0.103
	(1.000)	(0.869)	(0.738)	(0.923)	(0.782)	(1.000)	(0.900)	(0.617)	(0.969)
x_{4t}	1.603	-3.99e-04	2.81e-04	0.009	-0.056	0.010	0.958	-0.123	0.017
	(1.000)	(0.700)	(0.652)	(0.766)	(0.981)	(0.614)	(1.000)	(0.996)	(0.652)
x_{5t}	0.465	-6.38e-04	2.26e-04	0.021	-0.120	0.060	0.059	0.694	0.098
	(0.980)	(0.755)	(0.598)	(0.924)	(1.000)	(0.946)	(0.928)	(1.000)	(0.970)
x_{6t}	1.393	-2.49e-04	3.85e-04	0.002	0.077	-0.145	-0.088	0.081	0.754
	(1.000)	(0.619)	(0.685)	(0.562)	(0.996)	(1.000)	(0.994)	(0.938)	(1.000)
	$\lambda_0 \times 1200$		λ_x						
$\Delta s_t^{€}$	-1.469	0.019	0.204	-0.696	-0.275	-3.710	-0.230	0.274	-2.029
	(0.647)	(0.636)	(1.000)	(0.783)	(0.568)	(0.972)	(0.543)	(0.544)	(0.775)
$\Delta s_t^{£}$	-1.803	0.061	0.010	-0.869	0.510	-1.678	2.211	-4.172	1.589
	(0.678)	(0.856)	(0.575)	(0.826)	(0.614)	(0.795)	(0.841)	(0.919)	(0.714)
x_{1t}	0.093	2.19e-04	-1.84e-04	-0.029	-0.006	0.040	-0.019	-0.044	0.050
	(0.931)	(0.573)	(0.559)	(0.982)	(0.572)	(0.835)	(0.680)	(0.766)	(0.811)
x_{2t}	0.024	-3.83e-04	-2.85e-04	-0.025	0.044	-0.031	0.009	0.050	-0.051
	(0.673)	(0.619)	(0.593)	(0.984)	(0.950)	(0.836)	(0.607)	(0.860)	(0.880)
x_{3t}	-0.041	-3.20e-04	7.00e-04	-0.020	-0.037	0.027	0.065	-0.023	-0.013
	(0.736)	(0.581)	(0.679)	(0.910)	(0.865)	(0.746)	(0.946)	(0.649)	(0.600)
x_{4t}	0.004	-0.001	9.26e-04	0.015	-0.048	0.008	0.024	-0.114	0.022
	(0.524)	(0.825)	(0.796)	(0.895)	(0.965)	(0.592)	(0.761)	(0.991)	(0.693)
x_{5t}	-0.020	-3.21e-04	5.81e-04	0.019	-0.106	0.058	0.072	-0.234	0.080
	(0.621)	(0.590)	(0.665)	(0.909)	(1.000)	(0.937)	(0.964)	(1.000)	(0.935)
x_{6t}	0.063	-9.88e-04	2.56e-04	7.70e-04	0.090	-0.142	-0.103	0.089	-0.150
	(0.852)	(0.744)	(0.575)	(0.523)	(0.999)	(1.000)	(0.998)	(0.952)	(0.998)
	$\delta_{i,x}$		$\Sigma_x \times \sqrt{12} \times 100$						
$\Delta s_t^{€}$	0	9.728	0	0	0	0	0	0	0
	(—)	(1.000)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\Delta s_t^{£}$	0	5.311	8.714	0	0	0	0	0	0
	(—)	(1.000)	(1.000)	(—)	(—)	(—)	(—)	(—)	(—)
x_{1t}	1	-0.006	-0.018	0.157	0	0	0	0	0
	(—)	(0.769)	(0.989)	(1.000)	(—)	(—)	(—)	(—)	(—)
x_{2t}	0	-0.005	-0.015	0.096	0.076	0	0	0	0
	(—)	(0.788)	(0.991)	(1.000)	(1.000)	(—)	(—)	(—)	(—)
x_{3t}	0	-0.006	0.002	0.030	0.083	0.125	0	0	0
	(—)	(0.779)	(0.616)	(1.000)	(1.000)	(1.000)	(—)	(—)	(—)
x_{4t}	0	-0.006	2.59e-04	-0.041	-0.060	-0.015	0.096	0	0
	(—)	(0.842)	(0.513)	(1.000)	(1.000)	(0.999)	(1.000)	(—)	(—)
x_{5t}	0	-0.009	0.007	-0.062	-0.095	0.008	0.063	0.070	0
	(—)	(0.888)	(0.824)	(1.000)	(1.000)	(0.964)	(1.000)	(1.000)	(—)
x_{6t}	0	-0.006	-0.006	-0.036	-0.063	-0.084	0.041	0.011	0.068
	(—)	(0.786)	(0.816)	(1.000)	(1.000)	(1.000)	(1.000)	(0.999)	(1.000)

Posterior mean and probability that absolute value of the parameter is greater than zero (in parenthesis) of the WFX model. The state variables are: $x_{1t} = y_t^1$, $x_{2t} = y_t^1 - \widehat{y}_t^{€,1}$, $x_{3t} = y_t^1 - \widehat{y}_t^{£,1}$, $x_{4t} = y_t^{60} - y_t^1$, $x_{5t} = \left(y_t^{60} - y_t^1\right) - \left(\widehat{y}_t^{€,60} - \widehat{y}_t^{€,1}\right)$, $x_{6t} = \left(y_t^{60} - y_t^1\right) - \left(\widehat{y}_t^{£,60} - \widehat{y}_t^{£,1}\right)$