NBER WORKING PAPER SERIES

INTERNATIONAL YIELD CURVES AND CURRENCY PUZZLES

Mikhail Chernov
Drew D. Creal

Working Paper 25206
http://www.nber.org/papers/w25206

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2018, Revised February 2022

We thank the editor Stefan Nagel, the associate editor and three referees for their valuable feedback. We are also grateful to Michael Bauer, Magnus Dalhquist, Greg Duffee, Bernard Dumas, Federico Gavazzoni, Jeremy Graveline, Valentin Haddad, Sebastian Infante, Hanno Lustig, Philippe Mueller, Rob Richmond, Chris Telmer, Andrea Vedolin, Irina Zviadadze for comments on earlier drafts and participants in the seminars and conferences sponsored by Adam Smith Asset Pricing Conference 2019, Bundesbank Term Structure Workshop 2019, Federal Reserve Bank of San Francisco/Bank of Canada Conference on Fixed Income Markets 2019, McGill University, Santiago Finance Workshop 2018, SFS Cavalcade 2019, WU-ISK Symposium on Foreign Exchange Markets 2019. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

International Yield Curves and Currency Puzzles
Mikhail Chernov and Drew D. Creal
NBER Working Paper No. 25206
November 2018, Revised February 2022
JEL No. F31,G12,G15

ABSTRACT

The currency depreciation rate is often computed as the ratio of foreign to 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. This happens because of the FX bond disconnect, the inability of bonds to span exchange rates. Incorporating innovations to the pricing kernel that affect exchange rates but not bonds helps with resolving the puzzles. This approach also allows one to relate news about the cross-country differences between international yields to news about currency risk premiums.

Mikhail Chernov
Anderson School of Management
University of California, Los Angeles
110 Westwood Plaza, Suite C-417
Los Angeles, CA 90095
and CEPR
and also NBER
mikhail.chernov@anderson.ucla.edu

Drew D. Creal
University of Notre Dame
Department of Economics
dcreal@nd.edu

1 Introduction

A long line of research attempts to interpret exchange rate dynamics through the lens of pricing kernels. Backus, Foresi, and Telmer (2001) (B/F/T hereafter) show that with complete markets, currency depreciation rates equal the ratio of the two pricing kernels. Building on this observation, the asset market view (AMV) of exchange rates treats dynamic pricing kernels as primitives (Brandt, Cochrane, and Santa-Clara, 2006). Dynamic (i.e., conditional) pricing kernels are specified and estimated, with the goal of producing plausible exchange rate dynamics. The AMV approach has become the dominant paradigm in international financial economics.

The challenge in the AMV literature is matching observed exchange rate dynamics with those implied by the two estimated pricing kernels. Results hinge on both the functional forms for pricing kernels and the data used in estimation. Conditional pricing kernels are staples of the term structure literature, thus a natural first step towards implementing the AMV is to use domestic bond prices to estimate the domestic pricing kernel and foreign bond prices to estimate the foreign pricing kernel. Indeed, this is the approach pursued, explicitly or implicitly, by B/F/T, Bansal (1997), and many others.

While the specific details of implementation vary across papers, the key conclusion from this literature is that the depreciation rate inferred from the ratio of pricing kernels falls short of capturing the dynamics of depreciation rates that we observe in the data. A typical finding is that the variation in depreciation rates, inferred via the AMV, has little to do with 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). B/F/T conclude that modern term structure models cannot simultaneously fit the key properties of exchange rates while also fitting yields.

The more recent empirical literature attempts to both understand why the original paradigm fails and to produce models that are more successful at explaining FX dynamics. Since the AMV fails empirically, the proposed solutions attempt to depart from it. One branch rejects the AMV by pursuing an incomplete market approach that breaks the equivalence between depreciation rates and the ratio of pricing kernels. (e.g., Brandt and Santa-Clara, 2002). Another branch of this literature attempts to save the AMV framework by arguing that even if markets are incomplete one could still obtain depreciation rates via the the ratio of pricing kernels as long as those pricing kernels are estimated via variance minimization (e.g., Brennan and Xia, 2006, Sarno, Schneider, and Wagner, 2012). In practice, this literature takes us to the same approach as B/F/T as it infers pricing kernels using bond term structure models. Besides conceptual issues arising with both approaches, none of the proposed solutions is capable of capturing both the volatility and forward premium anomalies at once while also matching salient features of yields. Thus the challenge posed by B/F/T is still unresolved.

In this paper, we explain why the AMV fails empirically at inferring the depreciation rate when local currency bonds are used to estimate the respective pricing kernels. Then, we show how to successfully build a model that jointly matches features of depreciation rates and bond yields. Our model can resolve both the FX volatility and the FX forward premium anomalies, meeting the challenge posed by B/F/T. The key feature of a successful joint model of depreciation rates and yields is that the pricing kernel needs to include a shock to its martingale component that is uncorrelated with shocks to its transitory component.

We demonstrate that returns on local currency bonds do not span FX depreciation rates. We do so in our sample of five countries (Australia, Germany, Japan, UK, and USA). We regress FX returns on bond returns, both in levels and in logs. Further, as a robustness check, we consider samples before and after the Global Financial Crisis (GFC). The R^2 from these regressions are low. Therefore, one cannot use information in bonds alone to infer the dynamics of depreciation rates. Moreover, we can conclude that there must be shocks that impact depreciation rates that do not impact bonds.

Using this evidence, we build an international term structure model that simultaneously fits yields and depreciation rates. We take the perspective of a domestic (U.S.) investor who buys domestic and foreign bonds. To value foreign bonds, this investor discounts future cash flows on foreign bonds converted to domestic currency using their domestic pricing kernel. As a result, we abandon the AMV approach of estimating foreign and domestic pricing kernels and hoping their ratio equals observed depreciation rates. Nevertheless our approach allows computing the pricing kernel denominated in foreign currency via multiplication of the USD pricing kernel and the depreciation rate.

The domestic pricing kernel in our model is subject to two types of innovations. The first type impacts both bonds and depreciation rates while the second one impacts depreciation rates but not bonds. Our interpretation of the latter innovation is motivated by the literature on the multiplicative decomposition of the pricing kernel into the martingale and transitory components (Alvarez and Jermann, 2005, Hansen and Scheinkman, 2009). The element of the martingale component that is orthogonal to the transitory component does not affect bond prices (e.g., Mehra and Prescott, 1985). Therefore, we label the additional innovation in our model that impacts depreciation rates but not bonds (in combination with the convexity term associated with its exponential) as the purely martingale component of the pricing kernel.

We refer to this model as UFX (FX rates are unspanned by bonds). This model requires using data on US and foreign bond prices as well as exchange rates to identify all elements of the pricing kernel. This is in contrast to the approach that posits that FX is spanned by bonds (which we refer to as the SFX approach). An SFX model may be estimated with or without data on exchange rates. The latter is the predominant approach in the literature, and that is what we focus on in our analysis.

A lot of variations are possible within the affine framework. One could include macro variables (Brennan and Xia, 2006, Jotikasthira, Le, and Lundblad, 2015), UIP-based expectations of depreciation rates (Sarno, Schneider, and Wagner, 2012), or stochastic volatility (Anderson, Hammond, and Ramezani, 2010, Brandt and Santa-Clara, 2002). While all of these elements are important, they do not speak directly to our thesis about unspanned FX. Macroeconomic variables cannot help with spanning FX because of the macro disconnect puzzle. Expectations of depreciation rates cannot span depreciation rates themselves. FX volatility is, of course, time-varying and adding that feature is an obvious extension if one is interested in FX option valuation or other aspects of FX dynamics. But, on its own, it cannot span FX. We intentionally consider the simplest version of a term structure model that relies on bond yield principal components as driving variables so that we can highlight the advocated mechanism.

We show that both the UFX and SFX models match yields equally well. However, exchange rates implied by the SFX model are grossly misspecified. Its behavior is in line with findings reported by previous studies. Thus, the evidence is consistent with exchange rates that are unspanned by bonds.

In contrast, the UFX model implies realistic exchange rate behavior. The UFX model matches all the FX moments discussed by B/F/T and, in particular, both FX anomalies. We also consider evidence about the term structure of cross-sectional carry risk premiums from Lustig, Stathopolous, and Verdelhan (2019), known as the slope carry. That carry premiums decline with bonds' maturity is a natural complement to the B/F/T facts in the realm of FX-bond interaction. These authors argue that exchange rates should not feature the martingale component to explain the evidence. As we argue, such a property would imply the ability of local bonds to span depreciation rates. The UFX model is capable of matching the slope carry evidence despite featuring the martingale component in exchange rates.

The SFX model does not allow one to explore how currency risk premiums connect bond yields and bond risk premiums of different countries because it fails to capture their joint dynamics. We use the UFX model to interpret the differences between international yield curves. We decompose news about the currency risk premium at a given horizon into news about the expected future path of the depreciation rate and the cross-country bond yield differential. The latter contributes very little at short horizons with the contribution growing to about 50% at long horizons.

In our results, the quantitative impact of the purely martingale component is large. The innovations contribute 89 to 96 percent of the variation of depreciation rates, depending on the specific currency. The maximal Sharpe ratio associated with market prices of these innovations averages 0.3 on the monthly scale.

The flip side of using bonds to infer pricing kernels is an asset-pricing exercise in valuation of international bonds. 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.

The main lesson from our empirical study is that it is important to model a purely martingale component of a pricing kernel to capture FX dynamics. The presence of such a component suggests that a rich collection of international bonds does not span exchange rates. Following the literature on the FX macro disconnect, we refer to this finding as the *FX bond disconnect*. This lack of bond spanning should be a starting point for any equilibrium joint model of exchange rates and bond returns.

2 Background motivation

In this section we introduce definitions and notation used throughout the paper. After introducing these concepts, we revisit existing developments in the literature.

2.1 Bonds and currencies

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

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

where $G_{t,t+n}$ is the cash flow growth between time t and $t+n$. If the bond is issued in USD, then $G_{t,t+n} = 1$; 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 $G_{t,t+n} = S_{t+n}/S_t$ with S_t representing the nominal value of one unit of foreign currency in USD; its log is $s_t = \log S_t$. We denote the foreign bond price by Q_t^{*n} and its yield is y_t^{*n}. We use Δ to denote the time-series difference operator, e.g., $\Delta s_{t+1} = s_{t+1} - s_t$, and Δ_c to denote the US - other country difference operator, e.g., $\Delta_c y_t^{*n} = y_t^n - y_t^{*n}$.

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. A long-standing tradition in this literature is to specify dynamics of the pricing kernel $M_{t,t+n}$ expressed in USD and that same kernel expressed in foreign currency, $M^*_{t,t+n}$. The latter implies a value of a foreign-currency-denominated foreign-issued bond

$$Q_t^{*n} = E_t\left(M^*_{t,t+n}\right).$$

Next, researchers infer the depreciation rate via the ratio of the estimated pricing kernels

$$\frac{S_{t+1}}{S_t} = \frac{M^*_{t,t+1}}{M_{t,t+1}}. \tag{1}$$

Examples include, but are not limited to, Ahn (2004); Backus, Foresi, and Telmer (2001); Brennan and Xia (2006); Dahlquist and Hasseltoft (2013); Jotikasthira, Le, and Lundblad (2015); Kaminska, Meldrum, and Smith (2013); Sarno, Schneider, and Wagner (2012).

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

2.3 Complete markets and spanning

In this paper we argue that, in order to understand the described empirical evidence, one needs to explicitly account for local bonds' inability to span exchange rates. The difference between market completeness and this lack of spanning plays an important role in this paper. Thus, we define them here to clarify the distinction between the two.

2.3.1 Market completeness and the AMV

A market is complete if one can trade the full set of Arrow-Debreu securities, or has access to assets which can replicate these securities. This concept is pertinent for the justification of the relationship (1). As B/F/T discuss in detail, no-arbitrage implies that pricing kernels satisfy:

$$E_t\left(M^*_{t,t+1} \cdot R^*_{t+1}\right) = E_t\left(M_{t,t+1} \cdot \frac{S_{t+1}}{S_t} \cdot R^*_{t+1}\right), \tag{2}$$

where R^*_{t+1} is a foreign-currency denominated gross return on an asset. This relation holds if a pair of domestic and foreign pricing kernels satisfies equation (1). If markets are complete this pair is unique. Empirical implementation of this relation usually relies on the market completeness argument to justify taking the ratio of the two estimated pricing kernels, a.k.a. the AMV. Implicitly such an approach requires that the estimated pricing kernels match the true pricing kernels up to estimation noise. Put differently, bonds span exchange rates.

Here we connect to the literature that recognizes that financial markets are likely to be incomplete (e.g., Brennan and Xia, 2006, Sarno, Schneider, and Wagner, 2012). These authors argue that the AMV still holds if one instead considers minimum variance pricing kernels. This argument tries to tether a model-free line of reasoning that goes back to Hansen and Jagannathan (1991); Hansen and Richard (1987) to affine models. It fails in two dimensions.

First, the authors continue to estimate affine term structure models in which bonds span depreciation rates. In terms of actual implementation, their approach is essentially the same as B/F/T and Bansal (1997) and it is indistinguishable from the

AMV approach. Second, Burnside and Graveline (2019) and Sandulescu, Trojani, and Vedolin (2018) show that minimum variance pricing kernels do not allow the ratio of pricing kernels to reproduce the FX rate, despite the previous claim that they do.

Sandulescu, Trojani, and Vedolin (2018) establish that if one minimizes the pricing kernel's entropy (variance of the log pricing kernel in conditionally log normal models), then the ratio of the two pricing kernels recovers the exchange rate correctly even if markets are incomplete. Specifically, they project the USD pricing kernel on returns of both the U.S. and foreign assets. Their approach requires conversion of foreign returns into USD returns by using the observed exchange rate. Thus, the approach relies on exchange rate data as one of the inputs. The model that we advocate in the sequel, UFX, requires exchange rate data as an input to estimation as well. We estimate the log pricing kernel whose variance is the smallest out of all pricing kernels that match bond prices, similar to the entropy-based approach. The primary difference is that empirical implementation of the model-free projection-based view typically produces estimates of pricing kernels that price assets correctly only on average. Thus, this approach is not helpful to researchers who are interested in estimation of joint bond and FX dynamics.

2.3.2 Spanning regressions

Our thesis is that returns on local currency bonds do not span FX rates. As a result this feature has to be accommodated in a model. One can think of regressing returns (payoffs with prices equal to 1) of an asset of interest on returns (payoffs) of another set of assets. An $R^2 < 1$ implies that the former asset is unspanned by the latter.

That is what we do next in the context of local bonds spanning exchange rates.

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: $R_{t+1}^{FX} = S_{t+1}/S_t \times 1/Q_t^{*1}$. 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.

It is tempting to take returns on foreign bonds, $R_{t+1}^{*n} = Q_{t+1}^{*,n-1}/Q_t^{*n}$, convert them to USD returns, $S_{t+1}/S_t \times R_{t+1}^{*n}$, 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: $R_{t+1}^{*,FX} = S_t/S_{t+1} \times 1/Q_t^{*1}$, and regress these on foreign bond returns R_{t+1}^{*n}.

We work with monthly data from the US, UK, Australia, Japan, and Germany/Eurozone from January 1983 to April 2019 making for $T = 436$ 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 $12, 24, 36, 48, 60, 84$ and 120 months are downloaded from their respective central banks or government divisions (Federal Reserve, Bank of England, Japanese Ministry of Finance, Bundesbank, and the Reserve Bank of Australia).

That government bond yields are reliably available for maturities of one year and above is a typical issue in international finance. Thus, we follow a long tradition in the literature, and compute 1-month foreign interest rates by adding forward premiums to the US one-month yields (e.g., Dahlquist and Hasseltoft, 2016). One-period forward nominal exchange rates used for this computation are obtained from Datastream. US yields are from CRSP.[1]

The corresponding nominal exchange rates are from Datastream. Prior to the introduction of the Euro, we use the German Deutschemark and splice these series together beginning in 1999. Prior to 1987, only the 60 and 120 month yields are available for Australia. Prior to August 1986, the 120 month yield is missing for Japan.

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^{FX}_{t+12} on R^n_{t+12}, $n = 24, 36, 48, 60, \ldots, 120$. The R^2, regular and adjusted, from these regressions are reported in the top panel of Table 1, in the column labeled "\$ returns." The amount of variation in currency returns that can be hedged with bonds is quite modest, ranging between 15% for the British pound and 34% for the Japanese yen. The column labeled "FC returns" (FC stands for foreign currency) in the Table reports the R^2 from this regression, which are of a similar 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.

[1] This approach ignores the cross-currency Treasury basis which reflects the relative convenience yield (Jiang, Krishanmurthy, and Lustig, 2018). It is difficult to assess the quantitative effect of this omission in our case because these authors study the 12-month horizon due to the same data availability issues.

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 1. The resulting R^2 are similar to the ones without equities. The main asset classes are only weakly related to exchange rates.[2]

In order to better connect the evidence to affine models that we advocate in the paper, we also estimate the same regressions for log returns. We report the results in the bottom panel of Table 1. The evidence is very similar to that of gross returns.

Lastly, an important concern arising in the post Global Financial Crisis (GFC) world is whether bonds' inability to span exchange rates was affected by dramatic policy changes, which in particular led to near-zero interest rates throughout the world. To address this issue head-on we repeat the same regressions (gross returns only for brevity) in the pre- and post-GFC subsamples. Table 2 reports the results. We leave out the GFC itself by omitting the entire year of 2008.

Quantitatively, we observe an increase in the adjusted R^2 in both pre- and post-GFC periods as compared to the full sample. This is a manifestation of a structural change in the covariance structure of exchange rates and bond returns. Comparing the R^2 across the two subsamples, the post-GFC one features generally higher values, especially from the perspective of foreign currency returns. The largest values are for the Euro and British pound, which are close to 50%. Adding equity returns, increases the R^2 further, although not uniformly. For instance, FC returns for the Euro and

[2]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.

British pound are not affected, while the Japanese nearly double. Qualitatively, the evidence tells us the same story as the full sample: bonds cannot span exchange rates. We discuss below how to extend our advocated model to accommodate ultra-low interest rates and the zero lower bound.

These results imply that models whose innovations to bond returns also span depreciation rates will not be able to fit the two types of assets jointly. In the next section we advocate a modeling approach that breaks bond spanning.

3 Breaking the spanning of exchange rates with bonds

We first present a model that possesses features that are sufficient to achieve lack of spanning of exchange rates with bonds. Subsequently, we discuss the model's interpretation and possible extensions.

3.1 Model

We assume the N-dimensional state x_t follows a VAR(1):

$$x_t = \mu_x + \Phi_x x_{t-1} + \Sigma_x \varepsilon_t \qquad \varepsilon_t \sim \mathrm{N}(0, I), \tag{3}$$

It is important for the subsequent analysis that, under the null, the dimension N captures dynamics of all yields in their entirety. No other additional state variable helps improve the description of yields. The U.S. short interest rate is:

$$i_t = \delta_0 + \delta_x^\top x_t. \tag{4}$$

We model the dynamics of the pricing kernel expressed in USD as

$$-\log M_{t,t+1} = i_t + \frac{1}{2}\lambda_t^\top \lambda_t + \frac{1}{2}\gamma_t^\top \gamma_t + \lambda_t^\top \varepsilon_{t+1} + \gamma_t^\top \eta_{t+1}, \tag{5}$$

where λ_t are market prices of risk associated with shocks ε_t and γ_t is the market price of risk of the additional K-dimensional vector of shocks η_{t+1} that are in the pricing kernel.

The shocks η_{t+1} are central to breaking the spanning of depreciation rates with bonds in the model. First, they capture priced risks to which bonds are not exposed. Therefore, without loss of generality, we assume that η_{t+1} is independent of the vector ε_{t+1}. Second, since the short interest rate i_t is not exposed to η_t, USD-denominated credit-risk free bond yields will not be affected by the shock.

Indeed, the prices of U.S. zero-coupon bonds with maturity n are given by the standard pricing condition

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

Assuming market prices of ε-risk to be

$$\lambda_t = \lambda_0 + \lambda_x x_t, \tag{6}$$

US yields are linear functions of the factors x_t only

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

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

Third, bonds' inability to span depreciation rates is achieved by allowing the extra risk η_{t+1} to affect the latter. The K-dimensional vector of depreciation rates has dynamics

$$\Delta s_{t+1} \;=\; \mu_s + \Phi_{sx} x_t + \Sigma_{sx} \varepsilon_{t+1} + \Sigma_s \eta_{t+1}. \tag{7}$$

where we assume the dimension of the shock η_t equals the number of foreign countries in our sample. This is an oversimplification. One could explore multiple shocks per country, or, alternatively, impose some structure, e.g., common versus country-specific shocks. Our dataset with four foreign countries is not rich enough to distinguish between all these possibilities. The depreciation rate of country i is $\mathbf{e}_i^\top \Delta s_t$, where $\mathbf{e}_i$ is a unit vector. In the following, we avoid using country specific labels to conserve on notation and use Δs_t, or S_t/S_{t-1} in levels, when discussing specific countries.

We model the depreciation rate as opposed to the log exchange rate as a stationary variable in line with the large literature that views the log nominal exchange rate as being close to a unit root process. Anderson, Hammond, and Ramezani (2010) and Lustig, Stathopolous, and Verdelhan (2019) consider the alternative specification where the nominal exchange rate is stationary in log-levels. We revisit the importance of this assumption in light of the evidence on the term structure of currency carry risk premiums presented in the latter paper.

In our subsequent empirical work, we stay in the Gaussian framework and assume $\eta_t \sim \mathrm{N}(0, I)$ and the market price of η-risk is

$$\gamma_t = \gamma_0 + \gamma_x x_t \tag{8}$$

Then, equation (2) implies that the prices of zero-coupon foreign currency bonds with maturity n are given by

$$Q_t^{*n} = E_t\left(M_{t,t+1} \cdot \frac{S_{t+1}}{S_t} \cdot Q_{t+1}^{*n-1}\right). \tag{9}$$

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

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

where the bond loadings can be found in Appendix A.2. Thus, foreign credit-risk free bond yields are not affected by the shock η.

Since domestic and foreign bonds are only functions of x_t, the extra innovations $\Sigma_s \eta_t$ in (7) help match the variation in depreciation rates that are not captured by bonds. Also, because of this extra degree of freedom, the model's ability to capture realistic dynamics of the conditional mean of depreciation rates is not sacrificed. These dynamics are important for matching the forward premium as well as more recent empirical evidence on portfolios of currencies.

3.2 Discussion

3.2.1 Interpretation of the pricing kernel shock

This discussion naturally leads to the question of the economic nature of the shocks η. It is customary in the no-arbitrage literature to select shocks on the basis of assets that a model should explain. The vast majority of the international bond literature is an exception to this because of the literature's attempt to explain depreciation rates using bonds via the AMV. B/F/T is the prime example of this literature as they explicitly assume that bond returns span all of the shocks in the pricing kernel. Otherwise, there is nothing special in adding an extra shock relative to those that span bonds.

There is one important requirement for the shock: it cannot affect bonds precisely because bonds cannot span depreciation rates. There is nothing esoteric in such a requirement from an economic perspective. For example, an i.i.d. shock to consumption growth has a permanent effect on the marginal utility of the representative agent and does not impact bond prices (e.g., Alvarez and Jermann, 2005, Mehra and Prescott, 1985). More generally, this shock is related to the martingale component of the pricing kernel (Hansen and Scheinkman, 2009). We use this parallel to the equilibrium literature and refer to such shocks, in combination with the convexity term $\gamma_t^\top \gamma_t/2$, as the purely martingale component of the pricing kernel. The modifier "purely" emphasizes the element of the martingale component that is orthogonal to the transitory component of the pricing kernel.

Our specification is natural in the context of baseline international open-economy models with an exogenous stream of consumption in each country. For instance, the long-run-risk model of Bansal and Shaliastovich (2013) features country-specific shocks that in equilibrium affect bond returns and pricing kernels. By assuming complete financial markets, the authors derive the depreciation rate which becomes a function of both domestic and foreign shocks. Thus, bonds, which are subject to country-specific shocks only, are not able to span exchange rates. To assess this point quantitatively, we simulate 100,000 observations from this model economy and implement regressions of R^{FX}_{t+12} on R^{n}_{t+12}, and of R^{*FX}_{t+12} on R^{*n}_{t+12} $n = 12, 24, \ldots, 120$ that match the first column of Table 1. We have results for the UK only because that is what Bansal and Shaliastovich (2013) used to estimate the model in their paper. We find that the R^2 and adjusted R^2 are 4.54% and 4.53%, respectively. Thus, bonds cannot span currencies in this model.

Qualitatively, the same results may arise in 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. Unless the theoretical setting is such that bonds are capable of completing 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 the empirical evidence on the lack of consumption risk sharing. In such a setting, depending on the 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 thereof (e.g., Corsetti, Dedola, and Leduc, 2008). The latter result is consistent with our advocated framework.

3.2.2 Extensions

Depending on the intended application, the specification can be extended in a number of ways. For instance, one could contemplate a less restrictive version of the model. One could allow the depreciation rate in (7) to feedback on itself or enter into the conditional mean dynamics of x_t. In that case, one would need to impose additional restrictions as in Duffee (2011) to ensure that bond yields do not depend on the depreciation rate. There is little empirical evidence that depreciation rates predict themselves or interest rates, so this extension is likely to be more relevant in other settings. For completeness, we outline these restrictions in Appendix B.1.

Yet another interesting possibility to explore is whether the price of risk γ_t associated with the purely martingale component depends on the state of the macroeconomy. That entails making γ_t in (8) dependent on macro variables, such as output growth or economic uncertainty. To ensure that such a dependence is meaningful, that is, it cannot be spanned by yield factors, additional assumptions as in Joslin, Priebsch, and Singleton (2014) are needed. Appendix B.2 outlines such a specification.

Lastly, given the split-sample evidence on bond spanning presented earlier, one could capture the associated non-linearity by explicitly accounting for the zero lower bound (ZLB) interest rate environment post-GFC. For instance, Feunou, Fontaine, Le, and Lundblad (2021) build on the approach of Black (1995), Kim and Singleton (2012), and Xia and Wu (2016) by developing a ZLB extension of a Gaussian model in a way that allows for tractable estimation. Appendix Appendix B.3 sketches a model that satisfies the ZLB along these lines.

3.2.3 Connection to existing work

B/F/T and many authors following them explore so-called square-root, or CIR, state variables instead of the Gaussian ones used here, that is, making Σ_x time-varying. This distinction plays no role here. We are simply looking for models that are capable of a realistic fit to both domestic and foreign yield curves. 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 in the literature since then. A square-root factor can help capture time-varying volatility of interest rates, but, absent data on interest rate derivatives, it is hard to identify empirically (Bikbov and Chernov, 2011).

Jotikasthira, Le, and Lundblad (2015) regress depreciation rates on bond yields and conclude that an additional foreign exchange factor would be required. However, they do not specify a model with a set of factors or additional innovations that matches the FX anomalies. Specifically, they never state how the extra factor or innovation should appear in an affine model.[3]

Brandt and Santa-Clara (2002) feature equations similar to (3)-(5). However, they focus on modeling the wedge between the depreciation rate and the ratio of the two pricing kernels. That approach conflates bonds' inability to span exchange rates with market incompleteness. It leads the authors to introduce additional features such as stochastic volatility of depreciation rates, that is, time-varying Σ_s, and further

[3] To clarify, Jotikasthira, Le, and Lundblad (2015) model inflation and output as unspanned factors in their model. Their reason for doing so is different from ours. They use them to capture predictability of bond returns with these two macro variables rather than to model realistic FX dynamics. Inflation and output do not span FX, as evidenced by the FX macro disconnect puzzle.

shocks in order generate the associated time variation in variance. This structure requires complicated restrictions to ensure internal consistency (Burnside and Graveline, 2019, section 3.3). The paper does not explain the role of ε and η in the pricing kernel. "Bond spanning" or any idea like that is not mentioned.

Anderson, Hammond, and Ramezani (2010) do the reverse: they understand that market incompleteness is not central to the empirical success of modeling bonds and exchange rates jointly. However, the model they propose does not articulate the role of breaking bond spanning clearly. As a result, they follow Brandt and Santa-Clara (2002) by focusing on stochastic volatility of depreciation rates. As we explained above, while stochastic volatility could be a welcome modeling feature, it is not needed to generate lack of spanning by bonds in a model.

By modeling the price of all bonds from the perspective of the USD investor, we avoid using the AMV and do not have to take a stand on whether markets are complete or incomplete. This is because the valuation relationship (9) applies regardless of market completeness. We can focus exclusively on building a model without bond spanning of depreciation rates. We achieve that by introducing an extra vector of shocks η to the pricing kernel. One can recover FC-denominated pricing kernels from the USD denominated one and the corresponding depreciation rates. Furthermore, if markets are complete then such FC-denominated pricing kernels would match foreign pricing kernels.

That is the point of departure from the exploration of incomplete markets in affine settings studied by B/F/T and Lustig and Verdelhan (2019). These authors specify a model of the domestic and foreign pricing kernels, $M_{t,t+1}$ and $M^*_{t,t+1}$, respectively. Then they discuss conditions under which markets could be incomplete and how many assets are required to span the markets.

4 Results

4.1 Empirical approach

Our model should be able to jointly explain bonds of 8 maturities per each of the 5 countries and the corresponding 4 depreciation rates (versus the USD). Our model fits into the traditional state-space framework. As pointed out by Joslin, Singleton, and Zhu (2011), estimation of the model is simplified tremendously if one rotates the model into a set of observable factors that are assumed to be measured without error. That is the approach we implement here.

The choice of bond factors is motivated by the basic idea that three PCs explain almost all variation in the yield curve in a given country (Litterman and Scheinkman, 1991; Duffee, 2011). Next, we follow Sarno, Schneider, and Wagner (2012) and reduce the number of yield factors per country from three to two. This sacrifices the quality of fit to yields somewhat in exchange for tractability in a five-country setting. Lastly, we specify foreign variables as cross-country spreads over their U.S. counterparts, defined as U.S. variables minus foreign variables. Specifically, we select x_t to be

$$x_t = \begin{pmatrix} pc_t^{1,\$} \\ pc_t^{2,\$} \\ \Delta_c pc_t^{1,€} \\ \Delta_c pc_t^{2,€} \\ \Delta_c pc_t^{1,£} \\ \Delta_c pc_t^{2,£} \\ \Delta_c pc_t^{1,A\$} \\ \Delta_c pc_t^{2,A\$} \\ \Delta_c pc_t^{1,¥} \\ \Delta_c pc_t^{2,¥} \end{pmatrix} = \begin{pmatrix} \text{1st PC from US bonds} \\ \text{2nd PC from US bonds} \\ \text{1st PC from US vs German bonds} \\ \text{2nd PC from US vs German bonds} \\ \text{1st PC from US vs UK bonds} \\ \text{2nd PC from US vs UK bonds} \\ \text{1st PC from US vs Australian bonds} \\ \text{2nd PC from US vs Australian bonds} \\ \text{1st PC from US vs Japanese bonds} \\ \text{2nd PC from US vs Japanese bonds} \end{pmatrix}. \quad (10)$$

Our proposed model in Section 3.1 dictates that data on depreciation rates must be used during estimation. Otherwise, the effect of the purely martingale component is not econometrically identified. In order to distinguish shocks to the purely martingale component in Equation (7) from observation errors in a state-space model, we require the vector of depreciation rates, $(\Delta s_t^{€}, \Delta s_t^{£}, \Delta s_t^{A\$}, \Delta s_t^{¥})^\top$, to be fit perfectly. We refer to this model as the UFX model (unspanned foreign exchange rates).

We also estimate another model labeled as SFX (spanned foreign exchange rates) to establish a benchmark. In this model, we set $\eta_t = 0$ while keeping the dimension of x_t the same, $N = 10$. The SFX model does not require depreciation rates for estimation because, under the null, bond yields contain all the information about the shocks. While one could add depreciation rates as extra data in estimation, we choose not to do so. The primary reason is that most of the literature follows the same approach making it an appropriate benchmark. Another dimension of a typical implementation is to specify the domestic and foreign pricing kernels and then infer the depreciation rate via the AMV. Appendix B.4 connects our version of the bond spanning model to that approach.

Our estimation strategy has two ingredients. First, for both the UFX and SFX models, the state x_t is observable and is a linear transformation of yields. Since 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 (3). These parameters therefore require no identifying restrictions.

Second, we follow Joslin, Singleton, and Zhu (2011) and estimate risk premiums by specifying the risk-adjusted dynamics of the latent state that is rotated from the observed one so that the risk-adjusted persistence matrix of the latent state $\tilde{x}_t$ is diagonal. This dramatically simplifies identification of the risk-adjusted parameters as well as estimation. The extension to international yield curves is straightforward because depreciation rates are akin to macro variables in a single-country setting (e.g., Joslin, Priebsch, and Singleton, 2014). The mapping into risk-adjusted parameters and identifying restrictions are discussed in Appendix C.

In the UFX model, the dynamics of the depreciation rate, μ_s, Φ_{sx}, and Σ_{sx}, are identified directly from their dynamics in (7). In the SFX model, the parameters pertaining to the depreciation rate dynamics, that is, μ_s, Φ_{sx}, and Σ_{sx} are identified off the foreign yield curve. To see that, set Σ_s, γ_0 and γ_x to zero in the foreign bond loadings derived in Appendix A.2.

Given that we are jointly modeling the yield curve of five countries, the number of state variables is relatively large in order to fit the data. The number of parameters is consequently also large, especially for the VAR factor dynamics of the state x_t. Therefore, we impose two sets of overidentifying restrictions on the model.

We start with ex-ante restrictions. We assert that the persistence of the US factors is not affected by the foreign factors, while the persistence of foreign countries is affected by their own and US factors (e.g., the persistence of the UK factors is affected by the UK and US factors, but not by the German, Australian, or Japanese ones). These restrictions represent a US-centric view of the world, which is consistent with many international studies. Although we impose fewer restrictions, our approach is qualitatvely similar to that of Graveline and Joslin (2011).

We also impose ex-post restrictions on the parameters of Σ_x. We eliminate parameters that are insignificant across multiple countries at a time. Specifically, we treat each country in a symmetric fashion; a parameter is only restricted to zero in the persistence or covariance matrices if it can be restricted for all countries.

4.2 Properties of the estimated models

We start by discussing the models' implications for bond pricing in the first subsection. The remaining subsections are dedicated to various properties of the estimated exchange rate dynamics. In particular, we address the B/F/T challenge, quantitative importance of the unspanned shocks, and properties of the slope carry returns.

4.2.1 Estimates and fit to yields

We report the estimated parameters in Appendix D. Each model, SFX or UFX, has three groups of parameters pertaining to the VAR(1) dynamics of the state

x_t, dynamics of depreciation rates Δs_t, and "valuation". The last group contains the diagonal risk-adjusted persistence matrix of the rotated state $\tilde{x}_t$ and the factor loadings of interest rates and depreciation rates on that state. As $\tilde{x}_t$ corresponds to the latent factor rotation, the interpretation of the respective parameters is not intuitive. The main takeaway is that the valuation parameters are significant. Also, most of the VAR(1) parameters are significant.

Insignificant parameters are primarily associated with the properties of the depreciation rates in the UFX model. Our attempts to zero them out one-by-one and do so symmetrically across all countries led to counter intuitive results. This indicates the joint significance of these parameters, so we chose not to restrict them during estimation. Specifically, starting with the innovations η, almost all of the elements of Σ_s associated with the purely permanent component are significant. Elements of Σ_{sx} that reflect the effect of shocks to bonds on depreciation rates are primarily significant for the US shocks; this makes sense given our recursive identification of shocks that places the US first.

When we move to Φ_{sx}, which reflects the conditional expectation of depreciation rates, the US state variables are significant for AUD and JPY but insignificant otherwise. The remaining parameters that reflect the foreign yield factors ability to forecast depreciation rates are significant for AUD and JPY and insignificant for EUR and GBP. Amongst these, the loadings on the cross-country difference in the first PCs (e.g., $x_{3t} = \Delta_c pc_t^{1,€}, x_{5t} = \Delta_c pc_t^{1,£}, x_{7t} = \Delta_c pc_t^{1,A\$}, x_{9t} = \Delta_c pc_t^{1,¥}$) are of particular interest. These state variables are approximately equal to short interest rate differentials $\Delta_c y_t^{*1}$ and, thus, connect to the UIP regressions. All four of them are negative and significantly different from 1 consistent with UIP violations. Alternatively, one can relate their values to predictability of currency excess returns by subtracting 1. That version of the parameters is significantly less than zero for all four depreciation rates.

In contrast, many parameters comprising Φ_{sx} in the SFX model are significant. Without FX data, these parameters are identified off the foreign yield curve only. They essentially act as extra valuation parameters by expressing the pricing kernel in units of foreign currency. This happens without the discipline of matching exchange rates that are controlled by the same parameters. Thus, significance arises from the lack of this tension and the fit to a rich collection of foreign bonds.

We report the fit to yields along two dimensions. First, Table 3 displays pricing errors. Second, Table 4 compares the key moments of yields (mean, standard deviations, serial correlation) in the data and in the model. The overall message from this set of Tables is that both the SFX and UFX models fit the given collection of domestic and foreign yields similarly in terms of small errors, similar values of summary statistics, and values of estimated parameters that are related to yields.

4.2.2 Backus, Foresi, and Telmer

The 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 5 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 their similar fit to yields discussed earlier.

The differences in FX implications are drastic. While the UFX model matches the depreciation rate by construction, we need to infer them in the case of the SFX model. Given the estimated innovations to the yield factors x_t and depreciation rate parameters identified off foreign yield curves, we use Equation (7) to infer estimated realizations of depreciation rates. Figure 1 displays the inferred and observed depreciation rate. They are clearly different in terms of their scale and dynamic patterns.

According to Table 5, the inferred depreciation rate is 4 to 6 times more volatile in the SFX model than in the data, and the mean could be greater by two orders of magnitude (GBP or AUD), or have the wrong sign (JPY). The SFX model is nowhere close to the results of UIP regressions with positive and large slope coefficients. One way to interpret the SFX results is that bonds do not span FX rates, so it is incorrect to set the shock η to the purely martingale component to zero. Indeed, the UFX model replicates all of these moments perfectly, by construction. It does so without any sacrifice of the fit to yields, as discussed earlier.

To clarify, while the B/F/T methodology is close to the SFX model, it is not identical. B/F/T construct their model to match the UIP violations and the 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. This 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.

4.2.3 Variance decomposition and risk premiums

Moving beyond the baseline summary statistics highlighted by B/F/T, we investigate the quantitative importance of the unspanned shock η for variation in yields and depreciation rates. Further, we study the impact of prices of risk associated with this shock.

As our model nicely fits into the VAR framework, we can use a variance decomposition to quantify the impact of various innovations on both yield factors and depreciation rates. To minimize the impact of the order of variables that is required for such a decomposition, we organize innovations into 3 groups: US yields, non-US (foreign) yields, and depreciation rates. Table 6 presents the results.

The first two rows demonstrate the impact of the innovations ε and η on the US yield factors, x_{1t} and x_{2t}. By assumption, these factors are autonomous, so 100% of the variation is explained by the US innovations. The next eight rows display the impact on foreign yield factors. By construction, η has no impact on these either. US innovations have a large contribution to these factors (partially because the US is ordered first): the smallest impact of 55% is for $x_{8t} = \Delta_c pc_t^{2,A\$}$ and the largest impact of 92% is for $x_{5t} = \Delta_c pc_t^{1,£}$.

Moving on to the impact on the depreciation rates, which are ordered last, the smallest impact from the unspanned shock η is 89% for JPY. Thus, the shocks η are exceptionally important for capturing the risk of depreciation rates.

Does that translate into importance of risk premiums? Because η is orthogonal to ε, we can decompose the maximal Sharpe ratio (MSR) implied by our model (conditional volatility of the pricing kernel) into the contributions emanating from pricing the two types of shocks, $\sqrt{\lambda_t^\top \lambda_t}$ and $\sqrt{\gamma_t^\top \gamma_t}$, respectively. Figure 2 displays the time series of these MSRs.

The MSR associated with bond risks is consistent with the evidence reported elsewhere: the average is 0.78 with values ranging between 0.32 and 2.40. Overall, we observe a downward trend with elevated levels post-2008 and in the latter part of our sample. The MSR associated with exclusively currency risks averages 0.33. It is more stable than the bond MSR and significantly different from zero. It is statistically significantly lower than the bond MSR in the early part of the sample. Starting in the mid 1990s, the two are much closer to each other and their confidence bands often overlap. Thus, economically, the prices of risk associated with η are important.

Earlier, we demonstrated in split samples that bonds cannot span depreciation rates in both the pre- and post-GFC environment. Figure 2 suggests that the prices of risk from spanned and unspanned shocks became more closely aligned, although the alignment started pre-GFC in 2005. Furthermore, we observe statistically significant and economically large departures between the two in the immediate aftermath of 2008 and at the end of our sample.

4.2.4 The term structure of carry risk premiums

Lustig, Stathopolous, and Verdelhan (2019) establish a new property of depreciation rates that extends the B/F/T challenge about the FX-bond interaction. Specifically, they implement the cross-sectional carry one-month trading strategy replacing the traditional borrowing and lending that uses the one-month bond with bonds at longer maturities. Thus, the strategy is exposed to both FX and interest rate risks. Next, the authors achieve cross-sectional dispersion in returns by sorting currencies according to the slope of their domestic bond term structure; this strategy is known as the slope carry. This adds another maturity element to the approach. Lustig, Stathopolous, and Verdelhan (2019) demonstrate that cross-sectional carry returns decline with the maturity of the bonds used in the trading strategy.

We follow the same steps to establish the evidence in our sample. Our data set has a more limited cross-section of maturities (10 years versus 15 years). We use the slope as the sorting variable, defined as the difference between the 120-month and 12-month yields. The red dashed line and 95% confidence bands in Figure 3 display our estimates. It is qualitatively similar to Lustig, Stathopolous, and Verdelhan (2019) both in terms of point values and with substantive statistical uncertainty around them. Note that the dashed lines and the associated confidence bands are the same in both panels (A) and (B) of the Figure. They look different because of the scale.

Next, we construct the estimate from the UFX and SFX models. Specifically, we generate 25000 random simulated samples of length $T = 436$ from each model. For each sample, we implement the LSV cross-sectional sort to produce time-series of carry trades, by maturity. We compare the red dashed line to the average results across the 25000 samples from these Monte Carlo simulations (blue solid line). Also, we display the uncertainty bands of these results across the samples by plotting the the 2.5th and 97.5th percentiles.

The UFX model matches both the sample average and statistical uncertainty. That is not the case for the SFX model, as the 2.5th percentile from the data barely touches the 97.5th percentile from the model. The SFX-implied average returns deviate from the data by orders of magnitude, consistent with the B/F/T evidence of the previous section.

As the slope carry strategy is exposed to both yield and exchange rate fluctuations, a model's success in matching this pattern is contingent upon successfully capturing the joint distribution of these assets. We've demonstrated that the UFX model does a good job in this respect, and, while matching the new pattern is not necessarily guaranteed, this result is to be expected. In contrast, the SFX model does not capture the dynamics of depreciation rates, so its failure to capture the slope carry trade is not surprising.

A subtle point here is that if the average slope carry return is literally zero when using consol bonds as a source of funding, then the corresponding nominal exchange rate must be stationary (Backus, Boyarchenko, and Chernov, 2018, Lustig, Stathopolous, and Verdelhan, 2019). A stationary exchange rate does not have a martingale component at all (martingale components of the USD- and FC-denominated pricing kernel are identical). As we discussed throughout the paper, this property implies counterfactually that local bond returns can span depreciation rates.

So how is it possible that our model can replicate the evidence on the slope carry? In finite samples with finite maturities, it is difficult to discern if that return is zero indeed. In our sample where the longest maturity is 120 months, the average slope carry returns are significantly different from zero. In the Lustig, Stathopolous, and Verdelhan (2019) sample these returns cross zero at about 140 months and continue declining for longer maturities, but they are insignificant at any horizon. The issue is similar to testing a unit root versus the alternative of a highly persistent process, because in finite samples such processes behave similarly.

5 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 SFX model would be sufficient for that task. However, even a modest departure from this objective would render SFX-based analysis incomplete. This conclusion arises from relationships shared by international bonds and exchange rates. We first describe this relationship. Next, we provide examples of applications of yield-curve modeling that extend beyond matching yields.

5.1 Relation between bonds and currencies

Equation (2) combined with conditional log-normality implies

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

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

$$\Delta_c y_t^{*n} = es_t^n - srp_t^n + vs_t^n. \tag{11}$$

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.

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:

$$rx_{t+1}^n \equiv q_{t+1}^{n-1} - q_t^n + q_t^1 = -(n-1)y_{t+1}^{n-1} + ny_t^n - y_t^1 \tag{12}$$

with a similar expression for the foreign currency, rx_{t+1}^{*n}. Note that the reference rate for foreign excess returns is the short rate of the respective country, y_t^{*1}. Therefore, 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 (12) and (11) we get:

$$\begin{aligned} \Delta_c rx_{t+1}^{*n} &= (n-1) \cdot srp_{t+1}^{n-1} - n \cdot srp_t^n + srp_t^1 \\ &- (n-1) \cdot vs_{t+1}^{n-1} + n \cdot vs_t^n - vs_t^1 \\ &- u_{t+1}^n + u_{t+1}^1, \end{aligned} \tag{13}$$

where, for a given horizon j, $u_{t+1}^j = E_{t+1}\left(\sum_{i=1}^{j} \Delta s_{t+i}\right) - E_t\left(\sum_{i=1}^{j} \Delta s_{t+i}\right)$ – 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.

5.2 Applications of yield-curve modeling

Equations (11) and (13) demonstrate that researchers studying bond yields and bond risk premiums in the international context might be able to explain their properties by connecting them to currency risk premiums. The issue is that, while it would be natural to think that factors driving one type of risk premium should show up as factors driving the other, that is not the case because exchange rates are unspanned by bonds. There is at least one innovation that affects currency risk premiums but does not appear in bond risk premiums.

Therefore, the observed differences between bond yields or risk premiums, on their own, do not allow a researcher to identify the currency risk premiums, srp_t^n, 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.

The UFX model helps because it implies a realistic measure of srp_t^n by directly capturing the joint behavior of depreciation rates and prices of risk. We illustrate this in two different applications. First, we show that one would not be able to obtain the same result in an affine model that does not feature unspanned exchange rates, that is, in the SFX model. Second, we demonstrate that evidence about srp_t^n is useful for thinking about equilibrium models.

In the first application we ask whether it is still possible to decompose the yield differential, $\Delta_c y_t^{*n}$, into the currency risk premium and expected depreciation rate using the SFX model. As we have shown, the only way to use the SFX model itself is to use the AMV to infer an implicit depreciation rate and the corresponding risk premiums. That approach leads to a grossly misspecified depreciation rate (and non-nonsensical results for the decomposition of $\Delta_c y_t^{*n}$). To give the model the benefit of the doubt, an alternative approach would be to stick with the SFX implications for the joint distribution of innovations in yields and depreciation rates, but replace the SFX implications for srp_t^n with an established approach. We illustrate this strategy using the UIP regression to infer currency risk premiums.

In the second application, we implement this decomposition inside of an equilibrium model. This allows us to check whether that particular aspect of the relationship between cross-country differences in yields and currency risk premiums is captured by a structural model.

5.3 Decomposition of cross-country differences in yields

We start with equation (11) and, following Campbell (1991), define news, at horizon n, about differences in yields, the expected depreciation rate, and currency risk premium

$$
\begin{aligned}
N_{\Delta y,t}^n &\equiv \Delta_c y_t^{*n} - E_{t-1}\left(\Delta_c y_t^{*n}\right), \\
N_{s,t}^n &\equiv es_t^n - E_{t-1}\left(es_t^n\right), \\
N_{srp,t}^n &\equiv srp_t^n - E_{t-1}\left(srp_t^n\right).
\end{aligned}
$$

Then, equation (11) implies

$$\begin{aligned} N^n_{\Delta y,t} &= N^n_{s,t} - N^n_{srp,t}, \\ var(N^n_{\Delta y,t}) &= var(N^n_{s,t}) + var(N^n_{srp,t}) - 2cov(N^n_{s,t}, N^n_{srp,t}). \end{aligned} \tag{14}$$

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 (13). 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 news decomposition. 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 (14) around and work with:

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

The corresponding equation for variances is:

$$var(N^n_{srp,t}) \quad = \quad var(N^n_{s,t}) + var(N^n_{\Delta y,t}) - 2cov(N^n_{s,t}, N^n_{\Delta y,t}). \tag{15}$$

5.4 Decomposition results

To conserve space, we report our results for the British pound only in Figure 4. Results for other currencies are qualitatively similar and available in Appendix E.

We start with the UFX model, as a realistic benchmark. Given that the model is essentially a VAR, we compute the decomposition in equation (15) directly from the model. 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 the horizon grows, the share of news about the cross-country yield differential starts growing and reaches about 50% at the 2-year horizon. The relative importance of the two components equalizes at around 40 months. The news about the differences in yields dominate after that.

Next, as discussed above, we combine the SFX 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 difference between the two approaches is striking. First, the contribution of both es_t and $\Delta_c y_t^{*n}$ is greater than 100% at horizons below 15 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 UFX model. In the SFX model, the contribution of $\Delta_c y_t^{*n}$ is declining with horizon and it always stays above that of es_t.

Finally, we implement the same decomposition using the long-run-risk model of Bansal and Shaliastovich (2013). We have demonstrated that bonds cannot span exchange rates in this model, so it would be interesting to check how that relates to the decomposition of the currency risk premium. We use the formulas and estimated parameters reported by the authors to describe the joint evolution of exchange rates, yields, and state variables. Thus, we again obtain a VAR structure that allows us to compute the decomposition in equation (15).

Their paper estimates the US/UK model only at a quarterly frequency. We express quarterly horizons in months to facilitate the comparison with our estimated models. Finally, the authors report their theoretical UIP regressions in terms of real variables, while our evidence is based on nominal values. For robustness, we compute the decompositions from the model of Bansal and Shaliastovich (2013) for both real and nominal variables.

Both real and nominal economies are qualitatively similar to the results from the UFX model in that the contribution of $\Delta_c y_t^{*n}$ is small in the short term and increases with the horizon. In fact the real economy decomposition looks nearly identical to UFX. In the nominal case, the contribution of the difference in yields passes 100% at 60 months. Thus, despite a similar pattern, it misses the mark quantitatively.

6 Conclusion

We model the joint dynamics of international yield curves and exchange rates. In order to account for the associated risks, we develop a model where some of the innovations in currency depreciation rates do not affect bond yields. This feature of the model implies that innovations to bonds do not span exchange rates. To identify such innovations we combine estimation of yield curves with estimation of exchange rate dynamics. We find drastic differences in results relative to our benchmark model in which bonds span 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 large 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.

Alvarez, Fernando, and Urban Jermann, 2005, Using asset prices to measure the persistence of the marginal utility of wealth, *Econometrica* 73, 1977–2016.

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., Nina Boyarchenko, and Mikhail Chernov, 2018, Term structures of asset prices and returns, *Journal of Financial Economics* 129, 1–23.

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.

Bansal, Ravi, 1997, An exploration of the forward premium puzzle in currency markets, *The Review of Financial Studies* 10, 369–403.

———, and Ivan Shaliastovich, 2013, A long-run risks explanation of predictability puzzles in bond and currency markets, *The Review of Financial Studies* 26, 1–33.

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.

Black, Fischer, 1995, Interest rates as options, *Journal of Finance* 50, 1371–1376.

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.

Brandt, Michael W, John H. Cochrane, and Pedro Santa-Clara, 2006, International risk sharing is better than you think, or exchange rates are too smooth., *Journal of Monetary Economics* 53, 671–698.

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

Burnside, Craig, and Jeremy Graveline, 2019, On the asset market view of exchnage rates, *Review of Financial Studies.*

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.

Dahlquist, Magnus, and Henrik Hasseltoft, 2013, International bond risk premia, *Journal of International Economics* 90, 17–32.

——— , 2016, International bond risk premia., in P. Veronesi, ed.: *Handbook of Fixed-Income Securities* (Wiley).

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.

Feunou, Bruno, Jean-Sebastien Fontaine, Anh Le, and Christian Lundblad, 2021, Tractable term structure models, Management Science, forthcoming.

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.

Hansen, Lars, and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Hansen, Lars, and Scott Richard, 1987, The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models, *Econometrica* 55, 587–613.

Hansen, Lars Peter, and José A. Scheinkman, 2009, Long-term risk: An operator approach., *Econometrica* 77, 177–234.

Itskhoki, Oleg, and Dmitry Mukhin, 2020, Exchange rate disconnect in general eqilibrium, Journal of Political Economy, forthcoming.

Jiang, Zhengyang, Arvind Krishanmurthy, and Hanno Lustig, 2018, Foreign safe asset demand and the dollar exchange rate, Journal of Finance, forthcoming.

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 countries 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.

Kim, Don, and Kenneth Singleton, 2012, Term structure models and the zero bound: An empirical investigation of japanese yields, *Journal of Econometrics* 170, 32–49.

Litterman, Robert, and José Scheinkman, 1991, Common factors affecting bond returns, *The Journal of Fixed Income* 1, 54–61.

Lustig, Hanno, Andreas Stathopolous, and Adrien Verdelhan, 2019, The term structure of currency carry trade risk premia, *American Economic Review* 109, 4142–4177.

Lustig, Hanno, and Adrien Verdelhan, 2019, Does incomplete spanning in international financial markets help to explain exchange rates?, American Economic Review, forthcoming.

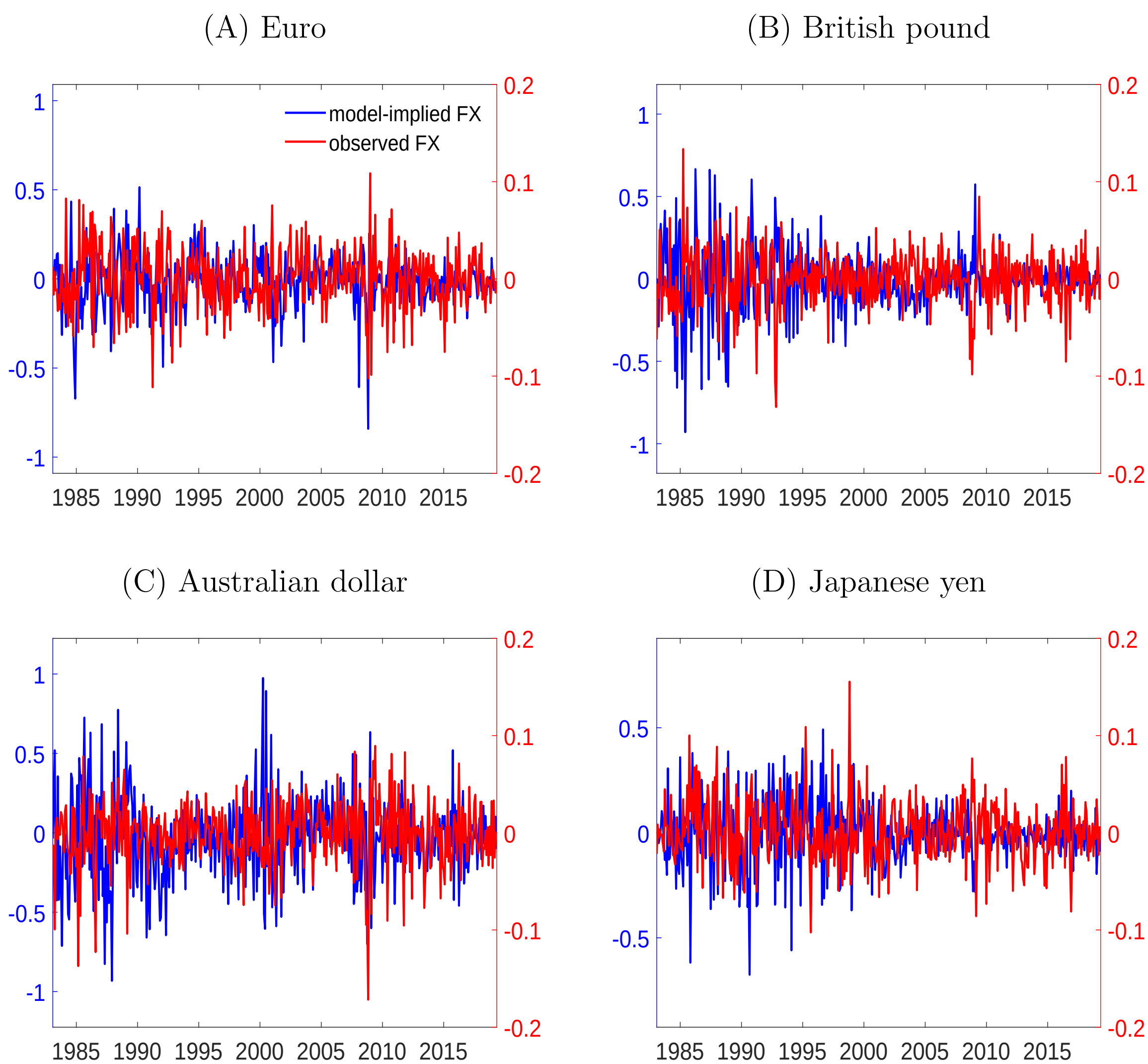
Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Sandulescu, Mirela, Fabio Trojani, and Andrea Vedolin, 2018, Model-free international stochastic discount factors, Journal of Finance, forthcoming.

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

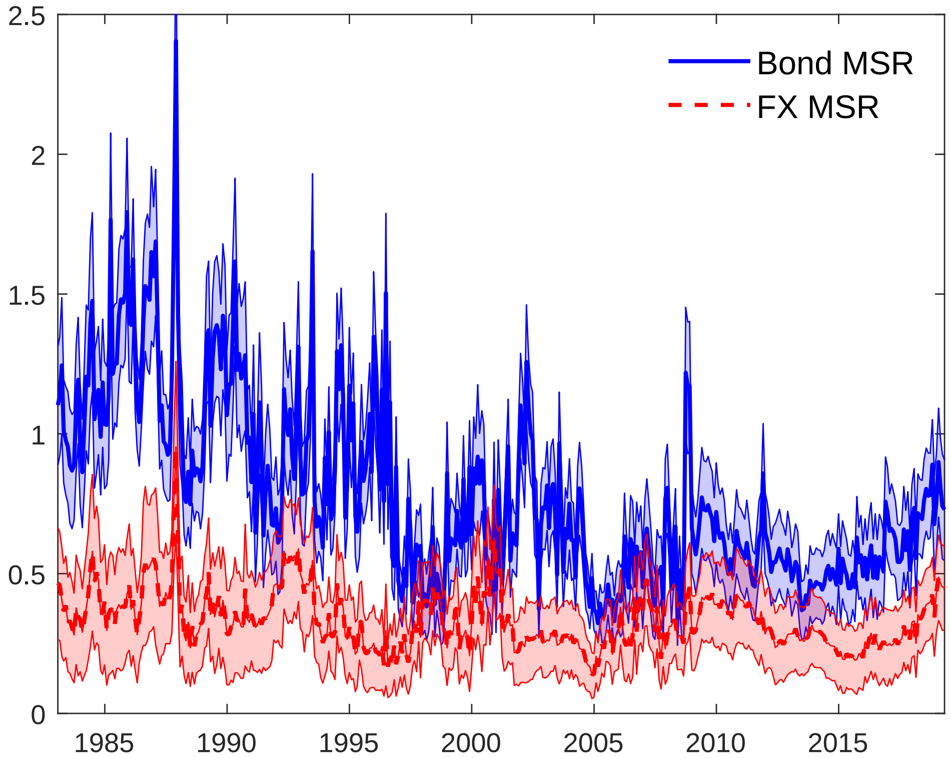
Xia, Fan Dora, and Jing Cynthia Wu, 2016, Measuring the macroeconomic impact of monetary policy at the zero lower bound, *Journal of Money, Credit, and Banking* 48, 253–291.

Figure 1
SFX-implied and observed depreciation rates



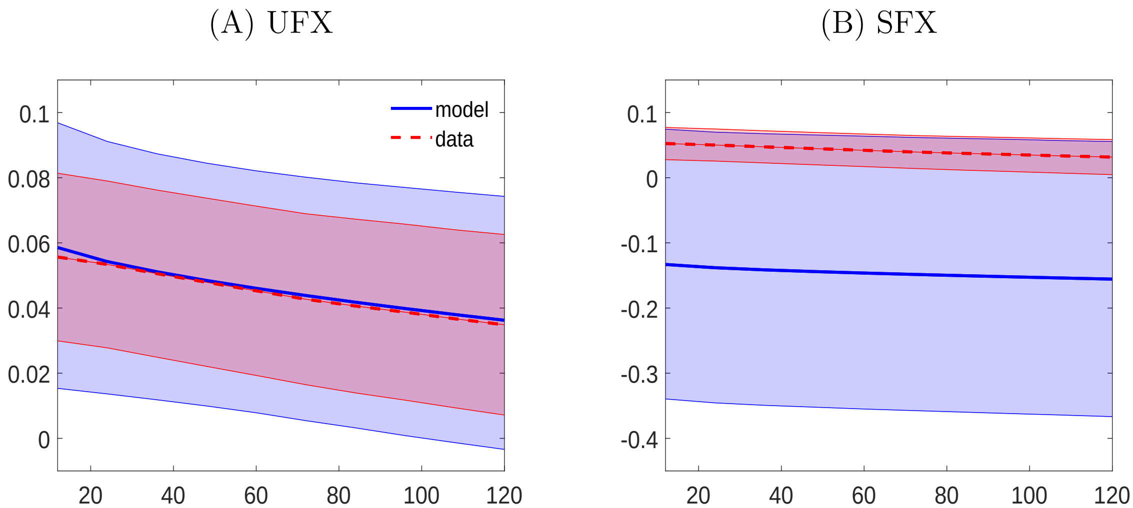
Notes: We plot the depreciation rates implied by the SFX model (blue, vertical axis left) against the observed depreciation rates (red, vertical axis right).

Figure 2
Maximal Sharpe ratios



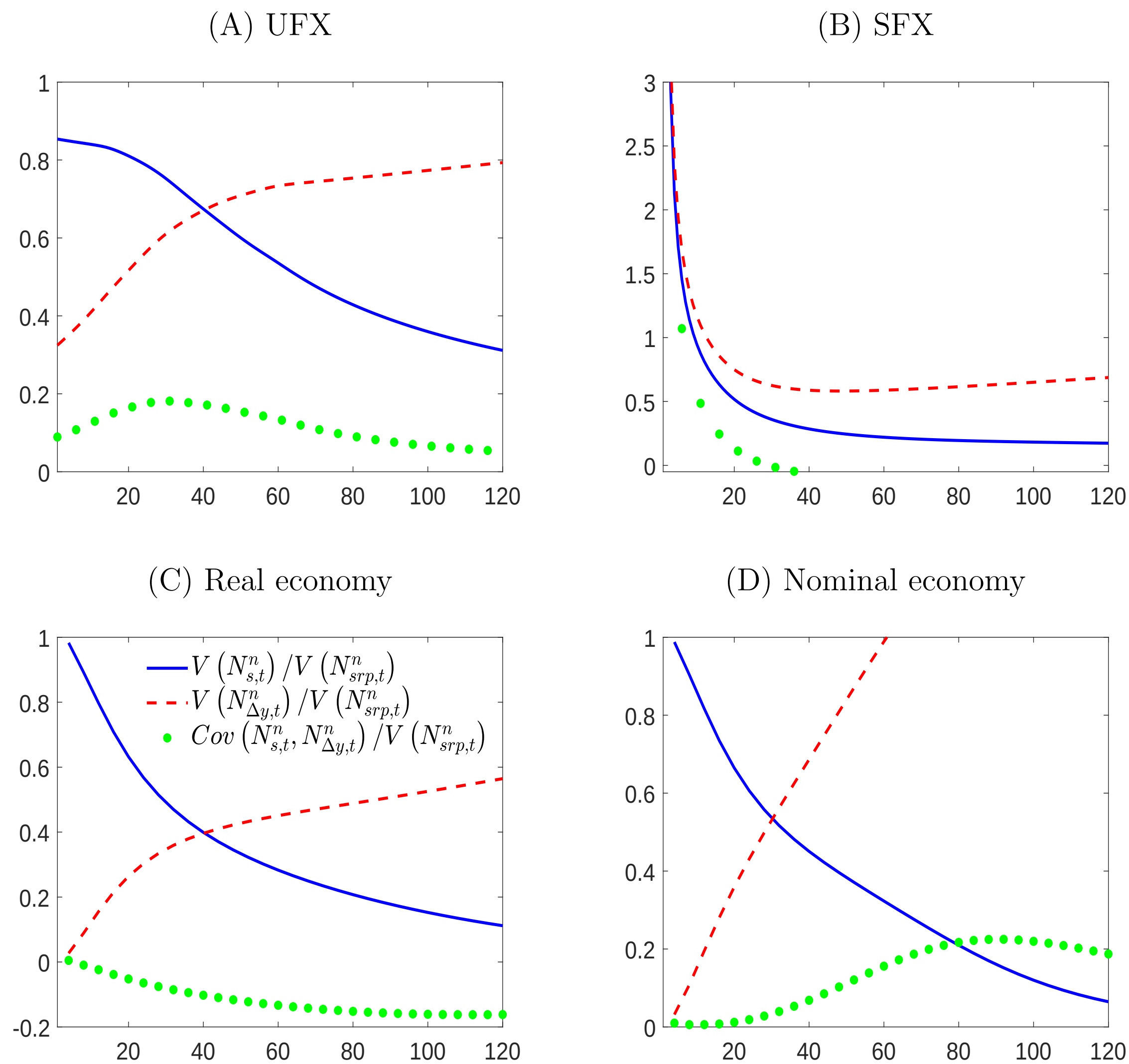
Notes: We decompose conditional volatility of the pricing kernel into the components associated with bond innovations ε, $\sqrt{\lambda_t^\top \lambda_t}$ *(blue line, Bond MSR), and depreciation rate innovations* η *that are unspanned by bonds,* $\sqrt{\gamma_t^\top \gamma_t}$ *(red line, FX MSR). The corresponding 95% confidence bands are depicted in lighter colors.*

Figure 3
Term structure of carry returns



Notes: We plot the unconditional average annual return of a cross-sectional carry trade as a function of maturity of the bonds that are used for borrowing and lending. The trading strategy uses the slope of the yield curve (120-month yield minus 12-month yield) as the sorting variable to create cross-sectional dispersion.

Figure 4
News-based decomposition of currency risk premiums, British pound



Notes: *We plot the percentage contribution to news about British pound 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*. We use three models to compute the decomposition: UFX, SFX, and real and nominal versions of the Bansal-Sahalistovich model.*

Table 1: Spanning regressions of currency returns on bond and equity returns

FX	Type of R^2	Bond returns		Bond and equity returns	
		\$ returns	FC returns	\$ returns	FC returns
Gross returns					
Euro	R^2	22.41	16.74	24.57	17.09
	R^2_{adj}	20.73	14.93	22.75	15.08
British pound	R^2	17.27	17.30	22.36	17.41
	R^2_{adj}	15.47	15.50	20.47	15.41
Australian dollar	R^2	21.44	24.45	25.84	26.53
	R^2_{adj}	19.50	22.59	23.80	24.52
Japanese yen	R^2	35.03	5.85	35.11	15.44
	R^2_{adj}	33.62	3.58	33.54	13.17
Log returns					
Euro	R^2	17.66	16.92	21.10	17.38
	R^2_{adj}	15.87	15.11	19.19	15.38
British pound	R^2	14.52	16.53	22.66	16.71
	R^2_{adj}	12.66	14.71	20.79	14.69
Australian dollar	R^2	23.02	25.25	27.79	26.99
	R^2_{adj}	21.12	23.40	25.81	24.99
Japanese yen	R^2	27.09	5.43	27.38	12.85
	R^2_{adj}	25.50	3.15	25.62	10.50

We report the R^2, regular and adjusted, expressed in percent for spanning regressions. We regress annual currency returns of a given country (obtained by investing in a foreign one-period bond) on annual bond returns of maturities $n = 2, 3, \ldots, 10$ years expressed in the same units (USD, denoted \$ returns, or foreign currency, denoted FC returns). We also combine bond returns with MSCI equity index returns in the last two columns.

Table 2: Spanning regressions of gross currency returns on bond and equity returns in sub samples

FX	Type of R^2	Bond returns		Bond and equity returns	
		$ returns	FC returns	$ returns	FC returns
January 1984 to December 2007					
Euro	R^2	31.59	14.74	31.60	16.94
	R^2_{adj}	29.38	11.98	29.13	13.94
British pound	R^2	13.96	16.75	14.47	20.83
	R^2_{adj}	11.16	14.05	11.37	17.96
Australian dollar	R^2	21.09	24.66	21.95	24.67
	R^2_{adj}	17.99	21.70	18.53	21.37
Japanese yen	R^2	38.89	13.81	38.98	13.81
	R^2_{adj}	36.91	10.52	36.78	10.15
January 2009 to April 2019					
Euro	R^2	26.91	50.26	39.75	51.84
	R^2_{adj}	21.14	46.33	34.41	47.58
British pound	R^2	29.04	56.88	53.27	56.90
	R^2_{adj}	23.44	53.48	49.14	53.08
Australian dollar	R^2	49.16	47.06	70.01	47.35
	R^2_{adj}	45.14	42.88	67.36	42.69
Japanese yen	R^2	40.68	35.64	40.84	75.62
	R^2_{adj}	36.00	30.56	35.60	73.47

We report the R^2, regular and adjusted, expressed in percent for spanning regressions. We regress annual currency returns of a given country (obtained by investing in a foreign one-period bond) on annual bond returns of maturities $n = 2, 3, \ldots, 10$ years expressed in the same units (USD, denoted $ returns, or foreign currency, denoted FC returns). We also combine bond returns with MSCI equity index returns in the last two columns.

Table 3: Yield pricing errors across countries

	SFX					UFX				
	USA	GER	UK	AU	JP	USA	GER	UK	AU	JP
y_t^{12}	0.06	0.11	0.10	0.11	0.11	0.06	0.11	0.10	0.11	0.11
y_t^{24}	0.06	0.08	0.07	0.06	0.06	0.06	0.08	0.07	0.06	0.06
y_t^{36}	0.05	0.05	0.05	0.04	0.04	0.05	0.05	0.04	0.04	0.04
y_t^{48}	0.04	0.02	0.03	0.02	0.03	0.04	0.02	0.03	0.02	0.03
y_t^{84}	0.03	0.05	0.04	0.04	0.06	0.03	0.05	0.04	0.04	0.06
y_t^{120}	0.08	0.09	0.09	0.07	0.08	0.08	0.09	0.08	0.07	0.07

Posterior mean estimates of the yield pricing errors in annualized percentage points for the US, Australia, Euro, Japan, and UK for both the SFX and UFX model, $100 \times [diag(\Sigma_y \Sigma_y^\top \times 12)]^{1/2}$. *Here* $\Sigma_y \Sigma_y^\top$ *is the covariance matrix of the measurement errors for yields of maturity* $n = 12, 24, 36, 48, 84, 120$.

Table 4: Sample moments of observed yields versus model-implied yields

	mean			st.dev.			autocorr		
	data	SFX	UFX	data	SFX	UFX	data	SFX	UFX
USA									
y_t^{12}	4.094	3.980	3.988	0.871	0.850	0.850	0.993	0.992	0.992
y_t^{24}	4.378	4.286	4.294	0.882	0.844	0.843	0.992	0.992	0.992
y_t^{36}	4.622	4.564	4.567	0.874	0.840	0.839	0.991	0.991	0.991
y_t^{48}	4.836	4.809	4.807	0.860	0.837	0.836	0.990	0.990	0.990
y_t^{84}	5.337	5.376	5.372	0.814	0.830	0.831	0.989	0.989	0.989
y_t^{120}	5.675	5.763	5.764	0.780	0.825	0.823	0.988	0.988	0.988
GER									
y_t^{*12}	3.419	3.192	3.207	0.794	0.775	0.773	0.993	0.992	0.992
y_t^{*24}	3.610	3.483	3.507	0.799	0.778	0.775	0.992	0.993	0.993
y_t^{*36}	3.806	3.742	3.760	0.799	0.782	0.780	0.992	0.993	0.993
y_t^{*48}	3.984	3.970	3.978	0.793	0.784	0.783	0.992	0.992	0.992
y_t^{*84}	4.399	4.491	4.473	0.759	0.777	0.778	0.991	0.991	0.990
y_t^{*120}	4.670	4.823	4.796	0.724	0.757	0.762	0.991	0.989	0.989
UK									
y_t^{*12}	5.288	5.159	5.181	1.116	1.094	1.096	0.984	0.985	0.984
y_t^{*24}	5.393	5.308	5.341	1.081	1.058	1.064	0.992	0.992	0.992
y_t^{*36}	5.506	5.451	5.480	1.051	1.035	1.041	0.992	0.991	0.991
y_t^{*48}	5.612	5.586	5.600	1.027	1.018	1.023	0.991	0.991	0.991
y_t^{*84}	5.854	5.904	5.883	0.973	0.983	0.979	0.990	0.990	0.990
y_t^{*120}	5.974	6.100	6.050	0.922	0.951	0.939	0.989	0.990	0.990
AU									
y_t^{*12}	5.773	5.613	5.655	0.996	0.974	0.980	0.985	0.985	0.984
y_t^{*24}	5.845	5.777	5.803	0.964	0.959	0.959	0.986	0.986	0.986
y_t^{*36}	5.972	5.949	5.963	0.947	0.951	0.948	0.986	0.987	0.987
y_t^{*48}	6.097	6.010	6.094	0.935	0.943	0.940	0.987	0.987	0.987
y_t^{*84}	6.356	6.405	6.385	0.912	0.913	0.913	0.988	0.987	0.987
y_t^{*120}	7.319	7.382	7.359	1.070	1.057	1.067	0.991	0.990	0.990
JP									
y_t^{*12}	1.812	1.537	1.570	0.721	0.706	0.709	0.991	0.986	0.987
y_t^{*24}	1.874	1.748	1.767	0.703	0.712	0.710	0.990	0.989	0.989
y_t^{*36}	1.993	1.942	1.950	0.706	0.715	0.710	0.990	0.989	0.990
y_t^{*48}	2.131	2.124	2.138	0.707	0.714	0.708	0.989	0.989	0.988
y_t^{*84}	2.520	2.591	2.555	0.709	0.695	0.694	0.989	0.988	0.988
y_t^{*120}	2.380	2.505	2.453	0.575	0.573	0.573	0.992	0.990	0.991

Sample moments from the yield data vs sample moments of model-implied yields. All yields have been annualized and multiplied by 100.

Table 5: Properties of Currency Prices and Interest Rates

	mean	SFX	UFX	st.dev.	SFX	UFX	autocorr	SFX	UFX
Panel A: Summary statistics									
	Δs_t								
Euro	0.899	-16.502	0.899	10.68	38.61	10.68	0.026	0.152	0.026
British pound	-0.595	-19.621	-0.595	10.09	69.00	10.09	0.051	-0.021	0.051
Australian dollar	-0.906	-27.959	0.906	11.49	77.23	11.49	0.054	0.018	0.054
Japanese yen	2.052	-7.674	2.052	10.81	46.95	10.81	0.046	-0.083	0.046
	Short rates								
USA	3.485	3.686	3.680	0.788	0.868	0.863	0.988	0.992	0.992
Germany	2.659	2.842	2.824	0.770	0.781	0.781	0.984	0.992	0.992
UK	4.861	5.000	4.969	1.134	1.160	1.156	0.991	0.994	0.994
Australia	6.142	6.266	6.253	1.182	1.195	1.198	0.981	0.986	0.986
Japan	0.952	1.291	1.193	0.685	0.684	0.675	0.963	0.988	0.983
	$y_t^1 - y_t^{*1}$								
Germany	0.722	0.734	0.746	0.670	0.692	0.688	0.984	0.989	0.989
UK	-1.480	-1.420	-1.399	0.628	0.604	0.605	0.974	0.979	0.979
Australia	-2.935	-2.878	-2.867	0.762	0.746	0.746	0.959	0.968	0.968
Japan	2.429	2.284	2.380	0.614	0.635	0.637	0.965	0.989	0.986

Panel B: UIP regressions

$$\Delta s_{t+1} = a + b\,(y_t^1 - y_t^{*1}) + \varepsilon_{t+1}$$

	$\widehat{a}$	SFX	UFX	$\widehat{b}$	SFX	UFX
Euro	0.0011	-0.0195	0.0011	-0.260	6.454	-0.270
	(0.0017)	(0.0058)	(0.0017)	(0.973)	(3.688)	(0.945)
British pound	-0.0014	-0.0170	-0.0015	-0.879	-2.582	-0.975
	(0.0015)	(0.0065)	(0.0015)	(1.095)	(5.457)	(1.156)
Australian dollar	-0.0020	-0.0066	-0.0022	-0.693	5.186	-0.770
	(0.0021)	(0.0134)	(0.0020)	(0.684)	(6.066)	(0.703)
Japanese yen	0.0044	-0.0003	0.0043	-1.316	-1.545	-1.277
	(0.0023)	(0.0082)	(0.0023)	(0.907)	(2.895)	(0.856)

In panel A we report the sample mean, standard deviation, and autocorrelation of the data. We compare them to the sample moments from model implied values from the UFX and SFX models. Depreciation rates, short rates, and interest rate differentials are annualized and multiplied by 100. In panel B we compare the UIP regression coeffcients in the data and in the two models.

Table 6: Variance decompositions: UFX model.

	US yields	foreign yields	FX
x_{1t}	1	0	0
x_{2t}	1	0	0
x_{3t}	0.70	0.30	0
x_{4t}	0.62	0.38	0
x_{5t}	0.44	0.56	0
x_{6t}	0.28	0.72	0
x_{7t}	0.46	0.54	0
x_{8t}	0.34	0.66	0
x_{9t}	0.43	0.57	0
x_{10t}	0.57	0.43	0
$\Delta s_t^{€}$	0.08	0.02	0.90
$\Delta s_t^{£}$	0.03	0.01	0.96
$\Delta s_t^{A\$}$	0.03	0.01	0.96
$\Delta s_t^{¥}$	0.08	0.03	0.89

The table displays how much variation in each variable is attributable to innovations in US yields, foreign yields, and depreciation rates. The numbers are obtained using a VAR-based variance decomposition. We group the yield of the four non-US countries together; the same goes for the depreciation rates.

Appendix A Bond prices

Appendix A.1 Domestic bonds in the UFX model

The U.S. short rate is

$$i_t = \delta_0 + \delta_x^\top x_t$$

The price of a one month bond is

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

which implies initial loadings $\bar{a}_1 = -\delta_0$ and $\bar{b}_{1,x} = -\delta_x$.

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

$$\begin{aligned}
Q_t^n &= E_t\left[M_{t,t+1} Q_{t+1}^{n-1}\right] \\
&= E_t\left[\exp\left(-\delta_0 - \delta_x^\top x_t - \frac{1}{2}\lambda_t^\top \lambda_t - \frac{1}{2}\gamma_t^\top \gamma_t - \lambda_t^\top \varepsilon_{t+1} - \gamma_t^\top \eta_{t+1} + \bar{a}_{n-1} + \bar{b}_{n-1,x}^\top x_{t+1}\right)\right] \\
&= \exp\left(\bar{a}_{n-1} - \delta_0 - \delta_x^\top x_t - \frac{1}{2}\lambda_t^\top \lambda_t - \frac{1}{2}\gamma_t^\top \gamma_t\right) E_t\left[\exp\left(-\lambda_t^\top \varepsilon_{t+1} - \gamma_t^\top \eta_{t+1} + \bar{b}_{n-1,x}^\top\left[\mu_x + \Phi_x x_t + \Sigma_x \varepsilon_{t+1}\right]\right)\right] \\
&= \exp\left(\bar{a}_{n-1} - \delta_0 - \delta_x^\top x_t + \bar{b}_{n-1,x}^\top\left[\mu_x + \Phi_x x_t\right] - \lambda_t^\top \Sigma_x^\top \bar{b}_{n-1,x} + \frac{1}{2}\bar{b}_{n-1,x}^\top \Sigma_x \Sigma_x^\top \bar{b}_{n-1,x}\right) \\
&= \exp\left(\bar{a}_{n-1} - \delta_0 - \delta_x^\top x_t + \bar{b}_{n-1,x}^\top\left[\mu_x + \Phi_x x_t\right] - \left(\lambda_0 + \lambda_x x_t\right)^\top \Sigma_x^\top \bar{b}_{n-1,x} + \frac{1}{2}\bar{b}_{n-1,x}^\top \Sigma_x \Sigma_x^\top \bar{b}_{n-1,x}\right) \\
&= \exp\left(\bar{a}_n + \bar{b}_{n,x}^\top x_t\right)
\end{aligned}$$

where the loadings are

$$\begin{aligned}
\bar{a}_n &= \bar{a}_{n-1} - \delta_0 + \bar{b}_{n-1,x}^\top\left(\mu_x - \Sigma_x \lambda_0\right) + \frac{1}{2}\bar{b}_{n-1,x}^\top \Sigma_x \Sigma_x^\top \bar{b}_{n-1,x} \\
\bar{b}_{n,x} &= \left(\Phi_x - \Sigma_x \lambda_x\right)^\top \bar{b}_{n-1,x} - \delta_x
\end{aligned}$$

Domestic 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 UFX model

The depreciation rate of country i has dynamics.

$$\Delta s_{t+1}^i = \mathbf{e}_i^\top \mu_s + \mathbf{e}_i^\top \Phi_{sx} x_t + \mathbf{e}_i^\top \Sigma_{sx} \varepsilon_{t+1} + \mathbf{e}_i^\top \Sigma_s \eta_{t+1}$$

where $\mathbf{e}_i$ is a unit vector that has a one in location i. In the following, we drop the country specific subscript. The price of a one month bond is

$$\begin{aligned}
Q_t^{*1} &= E_t\left[M_{t,t+1}\frac{S_{t+1}}{S_t}\right] \\
&= \exp\left(-\delta_0 - \delta_x^\top x_t - \frac{1}{2}\lambda_t^\top\lambda_t - \frac{1}{2}\gamma_t^\top\gamma_t + \mathbf{e}_i^\top\mu_s + \mathbf{e}_i^\top\Phi_{sx}x_t\right) \\
&\quad E_t\left[\exp\left(\left[\Sigma_{sx}^\top\mathbf{e}_i - \lambda_t\right]^\top\varepsilon_{t+1} + \left[\Sigma_s^\top\mathbf{e}_i - \gamma_t\right]^\top\eta_{t+1}\right)\right] \\
&= \exp\left(-\delta_0 - \delta_x^\top x_t - \frac{1}{2}\lambda_t^\top\lambda_t - \frac{1}{2}\gamma_t^\top\gamma_t + \delta_{s,0} + \mathbf{e}_i^\top\mu_s + \mathbf{e}_i^\top\Phi_{sx}x_t\right) \\
&\quad \exp\left(\frac{1}{2}\left(\Sigma_{sx}^\top\mathbf{e}_i - \lambda_t\right)^\top\left(\Sigma_{sx}^\top\mathbf{e}_i - \lambda_t\right) + \frac{1}{2}\left(\Sigma_s^\top\mathbf{e}_i - \gamma_t\right)^\top\left(\Sigma_s^\top\mathbf{e}_i - \gamma_t\right)\right) \\
&= \exp\left(-\delta_0 - \delta_x^\top x_t + \mathbf{e}_i^\top\mu_s + \mathbf{e}_i^\top\Phi_{sx}x_t + \frac{1}{2}\mathbf{e}_i^\top\Sigma_{sx}\Sigma_{sx}^\top\mathbf{e}_i - \lambda_t\Sigma_{sx}^\top\mathbf{e}_i + \frac{1}{2}\mathbf{e}_i^\top\Sigma_s\Sigma_s^\top\mathbf{e}_i - \gamma_t^\top\Sigma_s^\top\mathbf{e}_i\right) \\
&= \exp\left(-\delta_0 - \delta_x^\top x_t + \mathbf{e}_i^\top\mu_s + \mathbf{e}_i^\top\Phi_{sx}x_t + \frac{1}{2}\mathbf{e}_i^\top\Sigma_{sx}\Sigma_{sx}^\top\mathbf{e}_i + \frac{1}{2}\mathbf{e}_i^\top\Sigma_s\Sigma_s^\top\mathbf{e}_i\right) \\
&\quad \exp\left(-\left(\lambda_0 + \lambda_x x_t\right)\Sigma_{sx}^\top\mathbf{e}_i - \left(\gamma_0 + \gamma_x x_t\right)^\top\Sigma_s^\top\mathbf{e}_i\right) \\
&= \exp\left(\bar{a}_1^* + \bar{b}_{1,x}^{*\top}x_t\right)
\end{aligned}$$

where

$$\begin{aligned}
\bar{a}_1^* &= \delta_{s,0} - \delta_0 + \mathbf{e}_i^\top\left(\mu_s - \Sigma_{sx}\lambda_0 - \Sigma_s\gamma_0\right) + \frac{1}{2}\mathbf{e}_i^\top\left(\Sigma_{sx}\Sigma_{sx}^\top + \Sigma_s\Sigma_s^\top\right)\mathbf{e}_i \\
\bar{b}_{1,x}^* &= \left(\Phi_{sx} - \Sigma_{sx}\lambda_x - \Sigma_s\gamma_x\right)^\top\mathbf{e}_i - \delta_x
\end{aligned}$$

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

$$
\begin{aligned}
Q_t^{*n} &= E_t\left[M_{t,t+1}\frac{S_{t+1}}{S_t}Q_{t+1}^{*,n-1}\right]\\
&= \exp\left(\bar{a}_{n-1}^* - \delta_x^\top x_t - \frac{1}{2}\lambda_t^\top\lambda_t - \frac{1}{2}\gamma_t^\top\gamma_t + \mathbf{e}_i^\top\mu_s + \mathbf{e}_i^\top\Phi_{sx}x_t + \bar{b}_{n-1,x}^{*\top}\left[\mu_x + \Phi_x x_t\right]\right)\\
&\quad E_t\left[\exp\left(\left[\Sigma_{sx}^\top\mathbf{e}_i + \Sigma_x^\top\bar{b}_{n-1,x}^* - \lambda_t\right]^\top\varepsilon_{t+1} + \left[\Sigma_s^\top\mathbf{e}_i - \gamma_t\right]^\top\eta_{t+1}\right)\right]\\
&= \exp\left(\bar{a}_{n-1}^* - \delta_0 - \delta_x^\top x_t - \frac{1}{2}\lambda_t^\top\lambda_t - \frac{1}{2}\gamma_t^\top\gamma_t + \mathbf{e}_i^\top\mu_s + \mathbf{e}_i^\top\Phi_{sx}x_t + \bar{b}_{n-1,x}^{*\top}\left[\mu_x + \Phi_x x_t\right]\right)\\
&\quad \exp\left(\frac{1}{2}\left(\Sigma_{sx}^\top\mathbf{e}_i + \Sigma_x\widehat{\bar{b}}_{n-1,x} - \lambda_t\right)^\top\left(\Sigma_{sx}^\top\mathbf{e}_i + \Sigma_x^\top\bar{b}_{n-1,x}^* - \lambda_t\right) + \frac{1}{2}\left(\Sigma_s^\top\mathbf{e}_i - \gamma_t\right)^\top\left(\Sigma_s^\top\mathbf{e}_i - \gamma_t\right)\right)\\
&= \exp\left(\bar{a}_{n-1}^* - \delta_0 - \delta_x^\top x_t + \bar{b}_{n-1,x}^{*\top}\mu_x + \mathbf{e}_i^\top\mu_s + \bar{b}_{n-1,x}^{*\top}\Phi_x x_t + \mathbf{e}_i^\top\Phi_{sx}x_t\right)\\
&\quad \exp\left(\frac{1}{2}\bar{b}_{n-1,x}^{*\top}\Sigma_x\Sigma_x^\top\bar{b}_{n-1,x}^* + \bar{b}_{n-1,x}^{*\top}\Sigma_x\Sigma_{sx}^\top\mathbf{e}_i + \frac{1}{2}\mathbf{e}_i^\top\Sigma_{sx}\Sigma_{sx}^\top\mathbf{e}_i + \frac{1}{2}\mathbf{e}_i^\top\Sigma_s\Sigma_s^\top\mathbf{e}_i\right)\\
&\quad \exp\left(-\left(\lambda_0 + \lambda_x x_t\right)\Sigma_x^\top\bar{b}_{n-1,x}^* - \left(\lambda_0 + \lambda_x x_t\right)\Sigma_{sx}^\top\mathbf{e}_i - \left(\gamma_0 + \gamma_x x_t\right)^\top\Sigma_s^\top\mathbf{e}_i\right)\\
&= \exp\left(\widehat{\bar{a}}_n + \bar{b}_{n-1,x}^{*\top}x_t\right)
\end{aligned}
$$

where the loadings are

$$
\begin{aligned}
\bar{a}_n^* &= \bar{a}_{n-1}^* - \delta_0 + \bar{b}_{n-1,x}^{*\top}\left(\mu_x - \Sigma_x\lambda_0\right) + \mathbf{e}_i^\top\left(\mu_s - \Sigma_{sx}\lambda_0 - \Sigma_s\gamma_0\right) + \frac{1}{2}\bar{b}_{n-1,x}^{*\top}\Sigma_x\Sigma_x^\top\bar{b}_{n-1,x}^*\\
&\quad + \bar{b}_{n-1,x}^{*\top}\Sigma_x\Sigma_{sx}^\top\mathbf{e}_i + \frac{1}{2}\mathbf{e}_i^\top\left(\Sigma_{sx}\Sigma_{sx}^\top + \Sigma_s\Sigma_s^\top\right)\mathbf{e}_i\\
\bar{b}_{n,x}^* &= \left(\Phi_x - \Sigma_x\lambda_x\right)^\top\bar{b}_{n-1,x}^* + \left(\Phi_{sx} - \Sigma_{sx}\lambda_x - \Sigma_s\gamma_x\right)^\top\mathbf{e}_i - \delta_x
\end{aligned}
$$

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

Appendix B Extensions and alternative specifications

Appendix B.1 Unspanned factors in Gaussian models

We can consider more general dynamics for the domestic short rate, yield factors x_t and depreciation rates

$$
\begin{aligned}
i_t &= \delta_0 + \delta_x^\top x_t + \delta_s^\top\Delta s_t\\
x_t &= \mu_x + \Phi_x x_{t-1} + \Phi_{xs}\Delta s_{t-1} + \Sigma_x\varepsilon_t\\
\Delta s_t &= \mu_s + \Phi_{sx}x_{t-1} + \Phi_s\Delta s_{t-1} + \Sigma_{sx}\varepsilon_t + \Sigma_s\eta_t
\end{aligned}
$$

This model allows lagged values of the depreciation rate to predict itself or the yield factors. In addition, the market prices of risk can also depend on the depreciation rate.

$$
\begin{aligned}
\lambda_t &= \lambda_0 + \lambda_x x_t + \lambda_s \Delta s_t \\
\gamma_t &= \gamma_0 + \gamma_x x_t + \gamma_s \Delta s_t
\end{aligned}
$$

In order for the shock η_t to not influence yields, we need the following restrictions to hold

$$
\begin{aligned}
\delta_s &= 0, \\
\lambda_s &= \Phi_{xs}, \\
\gamma_s &= \Phi_s.
\end{aligned}
$$

If these restrictions hold, then foreign and domestic yields are linear functions of only the yield factors x_t and depreciation rates are unspanned meaning that both domestic and foreign yields are only a function of the yield factors

$$
\begin{aligned}
y_t^n &= a_n + b_n^\top x_t \\
y_t^{*n} &= a_n^* + b_n^{*\top} x_t
\end{aligned}
$$

The model in the main text is the special case when $\lambda_s = \gamma_s = 0$.

Appendix B.2 Spanned and unspanned macro factors

Another model extension is to add a vector of observable macro variables m_t to the model. The short rate, depreciation rate, and dynamics of the factors are

$$
\begin{aligned}
i_t &= \delta_0 + \delta_x^\top x_t + \delta_f^\top f_t + \delta_s^\top \Delta s_t \\
m_t &= \mu_m + \Phi_m m_{t-1} + \Phi_{mx} x_{t-1} + \Phi_{ms} \Delta s_{t-1} + \Sigma_m \varepsilon_t \\
x_t &= \mu_x + \Phi_{xm} m_{t-1} + \Phi_x x_{t-1} + \Phi_{xs} \Delta s_{t-1} + \Sigma_{xm} \varepsilon_t + \Sigma_x \varepsilon_t \\
\Delta s_t &= \mu_s + \Phi_{sm} m_{t-1} + \Phi_{sx} x_{t-1} + \Phi_s \Delta s_{t-1} + \Sigma_{sm} \varepsilon_t + \Sigma_{sx} \varepsilon_t + \Sigma_s \eta_t
\end{aligned}
$$

This model allows shocks to macro variables to contemporaneously impact exchange rates and yield factors. The dynamics of the conditional mean are also much more flexible for all variables. The log-stochastic discount factor in this model is

$$
-\log M_{t,t+1} = i_t + \frac{1}{2}\psi_t^\top \psi_t + \frac{1}{2}\lambda_t^\top \lambda_t + \frac{1}{2}\gamma_t^\top \gamma_t + \psi_t^\top \varepsilon_{t+1} + \lambda_t^\top \varepsilon_{t+1} + \gamma_t^\top \eta_{t+1}.
$$

The market prices of risk are

$$
\begin{aligned}
\psi_t &= \psi_0 + \psi_m m_t + \psi_x x_t + \psi_s \Delta s_t \\
\lambda_t &= \lambda_0 + \lambda_m m_t + \lambda_x x_t + \lambda_s \Delta s_t \\
\gamma_t &= \gamma_0 + \gamma_m m_t + \gamma_x x_t + \gamma_s \Delta s_t
\end{aligned}
$$

As in the previous section, we need restrictions to ensure that foreign and domestic yields are not functions of the depreciation rate.

$$\begin{aligned}\delta_s &= 0,\\ \psi_s &= \Phi_{ms},\\ \lambda_s &= \Phi_{xs},\\ \gamma_s &= \Phi_s.\end{aligned}$$

Yields are a function of the macroeconomic factors m_t.

Macroeconomic factors can either be spanned or unspanned at the same time that depreciation rates are unspanned. If a researcher wants macroeconomic factors to be unspanned as in Joslin, Priebsch, and Singleton (2014), then additional restrictions that can be imposed are

$$\begin{aligned}\delta_m &= 0,\\ \psi_m &= \Phi_m,\\ \lambda_m &= \Phi_{xm},\\ \gamma_m &= \Phi_m.\end{aligned}$$

Under these restrictions, yields are only a function of the yield factors x_t.

The interesting feature of this last setup is that one can test if risk premiums γ_t associated with the unspanned innovation in the depreciation rate η_t depend exclusively on the macro variables. Empirically, $\Phi_{ms} = 0$ (FX disconnect), $\Phi_{xs} = \Phi_s = 0$. The coefficient γ_x is ex-ante unrestricted, so one could test if that is equal to zero. If that is the case, γ_t is a function of factors m_t only.

Appendix B.3 Zero lower bound

The USD pricing kernel has exactly the same form as in Equation (5) with the same linear dependence of prices of risk on the state x_t. Following Black (1995), we change the functional form of the short interest rate to accommodate the zero lower bound:

$$i_t = \max(\delta_0 + \delta_x^\top x_t, 0).$$

The challenge in the international setting is to specify the depreciation rate so that the foreign short interest rate would not break the ZLB either, that is, we would like that rate to have the form

$$i_t^* = \max(\delta_0^* + \delta_x^{*\top} x_t, 0).$$

Thus the depreciation rate has to be such that it satisfies

$$i_t^* = -\log E_t\left(M_{t,t+1} \cdot \frac{S_{t+1}}{S_t}\right). \tag{B.1}$$

For simplicity, we specify a single depreciation rate. We guess its functional form as:

$$\Delta s_{t+1} = i_t - i_t^* + \widetilde{\mu}_s + \widetilde{\Phi}_{sx} x_t + \Sigma_{sx} \varepsilon_{t+1} + \Sigma_s \eta_{t+1}.$$

Compared to the main model in Equation (7) the innovations and innovation exposures (amount of risk) are identical. The conditional mean has changed. To emphasize that the parameters in this specification are different, we use a hat $\widehat{\cdot}$ over them. This functional form already assumes that the foreign interest rate stays above zero.

Next, we solve for the values of $\widetilde{\mu}_s$ and $\widetilde{\Phi}_{sx}$ that make Equation (B.1) hold. Substituting the dynamics for Δs_{t+1} and $m_{t,t+1}$ into (B.1) and solving, one obtains:

$$\begin{aligned} \widetilde{\mu}_s &= -\frac{1}{2}\left(\Sigma_{sx}\Sigma_{sx}^\top + \Sigma_s \Sigma_s^\top\right) + \lambda_0^\top \Sigma_{sx} + \gamma_0^\top \Sigma_s, \\ \widetilde{\Phi}_{sx} &= \lambda_x^\top \Sigma_{sx} + \gamma_x^\top \Sigma_s. \end{aligned}$$

Finally, one can use the approximation developed by Xia and Wu (2016) to value bonds of both domestic and foreign countries. As a last step, we can use the estimation methods for models at the ZLB developed by Feunou, Fontaine, Le, and Lundblad (2021), who extend the method of Joslin, Singleton, and Zhu (2011) to ZLB models.

Note that the ZLB affects the conditional mean of the depreciation rate via the interest rate differential, but not its volatility. This effect could be quite small, as, depending on parameter values,

$$i_t - i_t^* \approx \delta_0 + \delta_x^\top x_t - \delta_0^* - \delta_x^{*\top} x_t.$$

We leave careful analysis of these effects for future research.

Appendix B.4 Spanned FX model and the AMV approach

In the SFX model we posit the dynamics of the log pricing kernel expressed in USD as

$$-\log M_{t,t+1} = i_t + \frac{1}{2}\lambda_t^\top \lambda_t + \lambda_t^\top \varepsilon_{t+1}.$$

For simplicity, assume that a single depreciation rate has dynamics

$$\Delta s_{t+1} = \mu_s + \Phi_{sx} x_t + \Sigma_{sx} \varepsilon_{t+1}. \tag{B.2}$$

These two assumptions allow us to denominate the pricing kernel in foreign currency (FC) via

$$\begin{aligned} -\log M_{t,t+1}^* &= -\log M_{t,t+1} - \Delta s_{t+1} \\ &= -\mu_s - \Phi_{sx} x_t + i_t + \frac{1}{2}\lambda_t^\top \lambda_t + \left(\lambda_t^\top - \Sigma_{sx}\right)\varepsilon_{t+1}. \end{aligned}$$

Now, consider the AMV approach. The USD pricing kernel has the identical representation. Typically, the FC pricing kernel has a similar functional form:

$$
\begin{aligned}
-\log M^*_{t,t+1} &= i^*_t + \frac{1}{2}\lambda^{*\top}_t \lambda^*_t + \lambda^{*\top}_t \varepsilon_{t+1}, \qquad \text{(B.3)}\\
i^*_t &= \delta^*_0 + \delta^*_x x_t,\\
\lambda^*_t &= \lambda^*_0 + \lambda^*_x x_t.
\end{aligned}
$$

The AMV assumes market completeness and infers the depreciation rate via

$$
\begin{aligned}
\Delta s_{t+1} &= \log M^*_{t,t+1} - \log M_{t,t+1}\\
&= i_t - i^*_t + \frac{1}{2}\left(\lambda^\top_t \lambda_t - \lambda^{*\top}_t \lambda^*_t\right) + \left(\lambda_t - \lambda^*_t\right)^\top \varepsilon_{t+1}.
\end{aligned}
$$

If $\lambda_x = \lambda^*_x$, then the FC pricing kernel and inferred depreciation rate can be expressed as

$$
\begin{aligned}
-\log M^*_{t,t+1} &= \frac{1}{2}(\lambda_0 - \lambda^*_0)^\top (\lambda_0 - \lambda^*_0) - \lambda^\top_t (\lambda_0 - \lambda^*_0) + i^*_t + \frac{1}{2}\lambda^\top_t \lambda_t + \left(\lambda_t - [\lambda_0 - \lambda^*_0]\right)^\top \varepsilon_{t+1}, \qquad \text{(B.4)}\\
\Delta s_{t+1} &= \frac{1}{2}\left(\lambda^\top_0 \lambda_0 - \lambda^{*\top}_0 \lambda^*_0\right) + \left(\lambda_0 - \lambda^*_0\right)^\top \lambda_x x_t + i_t - i^*_t + \left(\lambda_0 - \lambda^*_0\right)^\top \varepsilon_{t+1}. \qquad \text{(B.5)}
\end{aligned}
$$

Comparing equations (B.3) and (B.4), or equations (B.2) and (B.5), we conclude that they have identical functional form. Thus, the two perspectives are empirically equivalent.

A natural question is whether removing the restriction $\lambda_x = \lambda^*_x$ can help the model fit both yields and depreciation rates without breaking the bond spanning. The previous version of this paper has implemented such a model and the answer is no. The restriction turns out to be helpful rather than hurting in a very specific sense. It helps combat parameter proliferation.

Appendix C Estimation

Appendix C.1 Observables

We stack the U.S. and foreign nominal yields of different maturities into vectors $y_t = \left(y^1_t, \ldots, y^{120}_t\right)^\top$ and $y^*_t = \left(y^{*1}_t, \ldots, y^{*120}_t\right)^\top$ as well as their bond loadings, e.g. $A = (a_1, \ldots, a_{120})^\top, B = (b_{1,x}, \ldots, b_{120,x})^\top$ and $A^* = (a^*_1, \ldots, a^*_{120})^\top, B^* = \left(b^*_{1,x}, \ldots, b^*_{120,x}\right)^\top$. For each country, the maturities include 1, 12, 24, 36, 48, 60, 84, and 120 months. In our data set, yields are missing for some countries and maturities.

When we include depreciation rates in the model, we write

$$\Delta s_t = \begin{pmatrix} \Delta s_t^{€} \\ \Delta s_t^{£} \\ \Delta s_t^{A\$} \\ \Delta s_t^{¥} \end{pmatrix}.$$

In our application, the vector Δs_t includes the depreciation rates from the Euro, U.K., Australia, and Japan.

We stack all the observables together in a vector Y_t. For the UFX model, the vector of observables is

$$Y_t = \begin{pmatrix} \Delta s_t \\ y_t \\ y_t^* \end{pmatrix}.$$

For the SFX model that does not include depreciation rates, this vector of observables is

$$Y_t = \begin{pmatrix} y_t \\ y_t^* \end{pmatrix}.$$

Appendix C.2 Augmented state vector

To implement the UFX model for estimation, it is easiest to write the dynamics in terms of an augmented state vector z_t that combines the yield factors x_t and the depreciation rates Δs_t. The state vector z_t contains only observable variables. We define z_t as

$$z_t = \begin{pmatrix} x_t \\ \Delta s_t \end{pmatrix}$$

The dynamics of the augmented state z_t are

$$\begin{aligned} z_t &= \mu_z + \Phi_x z_{t-1} + \Sigma_z w_t \\ &= \begin{pmatrix} \mu_x \\ \mu_s \end{pmatrix} + \begin{pmatrix} \Phi_x & 0 \\ \Phi_{sx} & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ \Delta s_{t-1} \end{pmatrix} + \begin{pmatrix} \Sigma_x & 0 \\ \Sigma_{sx} & \Sigma_s \end{pmatrix} \begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \end{aligned}$$

These dynamics form the basis of a linear, Gaussian state space form, which we discuss further below.

For the SFX model, the augmented state z_t and the yield factors x_t are the same since depreciation rates are not used in estimation. In this case, $z_t = x_t$.

For the UFX model, we can then write yields and depreciation rates as a linear function of the augmented state

$$\begin{aligned}\Delta s_t &= \omega_0 + \omega_z^\top z_t \\ y_t &= A + B_z z_t \\ y_t^* &= A^* + B_z^* z_t\end{aligned}$$

The vectors A and A^* contain the bond loadings as defined above. The remaining matrices are defined as

$$\begin{aligned}\omega_0 &= 0 \\ \omega_z^\top &= \begin{pmatrix} 0 & I \end{pmatrix} \\ B_z &= \begin{pmatrix} B & 0 \end{pmatrix} \\ B_z^* &= \begin{pmatrix} B^* & 0 \end{pmatrix}\end{aligned}$$

where B and B^* are the bond loadings for the U.S. and foreign countries.

Appendix C.3 Rotating the latent state vector to observables

Next, we define two matrices W_1 and W_2 that create linear combinations of the observables Y_t. Together, the concatenated matrix $\left(W_1^\top; W_2^\top\right)^\top$ 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 a linear combination (of yields) observed with error.

Following the term structure literature, the matrix W_1 is chosen so that the augmented state vector z_t is a linear combination of observables

$$z_t = \begin{pmatrix} x_t \\ \Delta s_t \end{pmatrix} = Y_t^{(1)} = W_1 Y_t$$

with x_t defined as in the text for both the UFX and SFX models. For the SFX model, the matrix W_1 includes eigenvectors of the sample covariance matrix associated with the first two principal components for each country. For the UFX model, the matrix W_1 includes four rows of the unit vector that select out of Y_t the depreciations rates Δs_t as well as the eigenvectors of the sample covariance matrix associated with first two principal components for each country.

Let q_j denote the j-th eigenvector of the sample covariance matrix for yields across all 8 maturities for a given country. We use a superscript to denote a country, e.g. $q_j^{\$}$ is an eigenvector for the

United States. The matrix W_1 written in terms of blocks of the first two eigenvectors for each country is

$$W_1 = \begin{pmatrix} 0 & q_1^{\$,\top} & 0 & 0 & 0 & 0 \\ 0 & q_2^{\$,\top} & 0 & 0 & 0 & 0 \\ 0 & q_1^{\$,\top} & -q_1^{€,\top} & 0 & 0 & 0 \\ 0 & q_2^{\$,\top} & -q_2^{€,\top} & 0 & 0 & 0 \\ 0 & q_1^{\$,\top} & 0 & -q_1^{£,\top} & 0 & 0 \\ 0 & q_2^{\$,\top} & 0 & -q_2^{£,\top} & 0 & 0 \\ 0 & q_1^{\$,\top} & 0 & 0 & -q_1^{A\$,\top} & 0 \\ 0 & q_2^{\$,\top} & 0 & 0 & -q_2^{A\$,\top} & 0 \\ 0 & q_1^{\$,\top} & 0 & 0 & 0 & -q_1^{¥,\top} \\ 0 & q_2^{\$,\top} & 0 & 0 & 0 & -q_2^{¥,\top} \\ I_4 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The identity matrix I_4 in the last row selects out the depreciation rates, which are ordered first in Y_t.

The matrix W_2 are linear combinations of Y_t that are not used when defining z_t. Specifically, in our work, we choose W_2 to equal the remaining eigenvectors of the sample covariance matrix associated with the third through eighth principal components from each country. Recall that the first two principal components are used to define W_1.

In this appendix, we use a 'tilde' to denote any parameters $\tilde{\theta}$, factor loadings $\tilde{B}$, or state variables $\tilde{x}_t$ (or $\tilde{z}_t$) under the latent factor rotation. Consider the UFX model. The augmented state vector $\tilde{z}_t$ is defined as

$$\tilde{z}_t = \begin{pmatrix} \tilde{x}_{1t} \\ \tilde{x}_{2t} \end{pmatrix}$$

For the UFX model, this is a $N + M \times 1$ vector of latent factors.

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

$$Y_t = \begin{pmatrix} \Delta s_t \\ y_t \\ y_t^* \end{pmatrix} = \begin{pmatrix} \omega_0 \\ \tilde{A} \\ \tilde{A}^* \end{pmatrix} + \begin{pmatrix} \omega_z^\top \\ \tilde{B}_z \\ \tilde{B}_z^* \end{pmatrix} \tilde{z}_t = \tilde{C} + \tilde{D}\tilde{z}_t$$

where the vector $\tilde{C}$ and matrix $\tilde{D}$ contain the stacked intercepts and factor loadings under the latent factor rotation. The observed factors z_t are related to the latent factors $\tilde{z}_t$ through the linear transformation

$$z_t = \Gamma_0 + \Gamma_1 \tilde{z}_t. \tag{C.6}$$

Next, we pre-multiply Y_t above by W_1 and substitute out the latent state vector $\tilde{z}_t = \Gamma_1^{-1}(z_t - \Gamma_0)$

for the observed state

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

To guarantee that the observed state is $z_t = W_1 Y_t$, the rotation matrices Γ_0 and Γ_1 in (C.6) must satisfy the restrictions

$$
\Gamma_0 = W_1 \tilde{C} \tag{C.7}
$$

$$
\Gamma_1 = W_1 \tilde{D} \tag{C.8}
$$

for any parameters θ in the model. Moreover, the matrix Γ_1 must be invertible.

Appendix C.4 Parameterization and identification

The risk premium 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-adjusted 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.

Appendix C.4.1 UFX model

In the appendix, we use Q to denote parameters under the risk neutral measure. The risk neutral dynamics under the latent factor rotation are

$$
\begin{aligned}
\Delta s_t &= \tilde{\omega}_0 + \tilde{\omega}_z^\top \tilde{z}_t \\
i_t &= \tilde{\delta}_0 + \tilde{\delta}_z^\top \tilde{z}_t \\
\tilde{z}_t &= \tilde{\mu}_z^Q + \tilde{\Phi}_z^Q \tilde{z}_{t-1} + \tilde{\Sigma}_z w_t^Q
\end{aligned}
$$

In this model, there is only one set of risk-adjusted parameters.

Under this rotation, the model is identified by imposing the following restrictions on the factor loadings for the U.S. short rate and the depreciation rates

$$
\begin{pmatrix} \tilde{\delta}_z & \tilde{\omega}_z^{€} & \tilde{\omega}_z^{£} & \tilde{\omega}_z^{A\$} & \tilde{\omega}_z^{¥} \end{pmatrix} = \begin{pmatrix}
1 & \tilde{\omega}_1^{€} & \tilde{\omega}_1^{£} & \tilde{\omega}_1^{A\$} & \tilde{\omega}_1^{¥} \\
1 & \tilde{\omega}_2^{€} & \tilde{\omega}_2^{£} & \tilde{\omega}_2^{A\$} & \tilde{\omega}_2^{¥} \\
\tilde{\delta}_3 & 1 & \tilde{\omega}_3^{£} & \tilde{\omega}_3^{A\$} & \tilde{\omega}_3^{¥} \\
\tilde{\delta}_4 & 1 & \tilde{\omega}_4^{£} & \tilde{\omega}_4^{A\$} & \tilde{\omega}_4^{¥} \\
\tilde{\delta}_5 & \tilde{\omega}_5^{€} & 1 & \tilde{\omega}_5^{A\$} & \tilde{\omega}_5^{¥} \\
\tilde{\delta}_6 & \tilde{\omega}_6^{€} & 1 & \tilde{\omega}_6^{A\$} & \tilde{\omega}_6^{¥} \\
\tilde{\delta}_7 & \tilde{\omega}_7^{€} & \tilde{\omega}_7^{£} & 1 & \tilde{\omega}_8^{¥} \\
\tilde{\delta}_8 & \tilde{\omega}_8^{€} & \tilde{\omega}_8^{£} & 1 & \tilde{\omega}_8^{¥} \\
\tilde{\delta}_9 & \tilde{\omega}_9^{€} & \tilde{\omega}_9^{£} & \tilde{\omega}_9^{A\$} & 1 \\
\tilde{\delta}_{10} & \tilde{\omega}_{10}^{€} & \tilde{\omega}_{10}^{£} & \tilde{\omega}_{10}^{A\$} & 1 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix} \tag{C.9}
$$

The number of ones in each column is consistent with the matrix W_1 used to define the observable factors. For example, two observable factors are the first two U.S. principal components, which implies two ones in $\tilde{\delta}_z$. For each foreign country, there are three observable factors. These include the depreciation rate and the first two principal components (minus the first two PC's from the United States). Therefore, each column $\tilde{\omega}_z$ needs three restrictions.

To identify the factors' location, the drifts are set to zero under the latent factor rotation

$$
\tilde{\mu}_z^Q = 0
$$

The intercepts $\tilde{\delta}_0$ and $\tilde{\omega}_0^i$ for each country i are estimable.

Finally, we assume that the risk-adjusted autocovariance matrix is a diagonal matrix of eigenvalues.

$$
\tilde{\Phi}_z^Q = \begin{pmatrix} \tilde{\Phi}_x^Q & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix}
\tilde{\phi}_1^Q & 0 & 0 & \dots & 0 & 0 \\
0 & \tilde{\phi}_2^Q & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & \dots & \dots & \tilde{\phi}_N^Q & \dots & 0 \\
\vdots & \vdots & \dots & \dots & & \vdots \\
0 & \dots & \dots & \dots & \ddots & \vdots \\
0 & \dots & \dots & \dots & \dots & 0
\end{pmatrix}
$$

The matrix $\tilde{\Phi}_z^Q$ has dimension $N + M \times N + M$. We emphasize that the last M eigenvalues in this matrix are zero (the lower right block). In general, the matrix $\tilde{\Phi}_z^Q$ is required to be a matrix of eigenvalues to streamline identification. 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, which is consistent with distinct, real eigenvalues ordered from largest to smallest.

Appendix C.4.2 SFX model

In the SFX model, the identifying restrictions are exactly the same as the UFX model. In the SFX model, we set the matrix $\Sigma_s = 0$.

Appendix C.5 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_z is lower triangular. For the UFX model, this matrix has diagonal elements σ_s and σ_x, associated with the depreciation rates and yield factors. For the SFX model, the matrix contains only the values σ_x associated with the yield factors. For the depreciation rates, we place an inverse Gamma prior distribution on the variances $\sigma_{s,j}^2 \sim \text{IG}\left(\alpha_s, \beta_s\right)$ for $j = 1, \ldots, M$. For the yield factors, we place an inverse Gamma prior distribution on the variances $\sigma_{x,j}^2 \sim \text{IG}\left(\alpha_x, \beta_x\right)$ for $j = 1, \ldots, N$. We calculate the OLS estimate of Σ_z from the VAR and choose the hyperparameters $\{\alpha_s, \beta_s, \alpha_x, \beta_x\}$ to match the first moment and have a large variance.
- The location parameters in the model include the unconditional means $\bar{\mu}_z$ of the VAR factor dynamics and the drifts $\tilde{\delta}_0$ and $\tilde{\omega}_0^i$ for each country i under the latent factor rotation. First, we calculate the unconditional sample mean of the factors $\hat{\bar{\mu}}_z$. Our prior for each element of $\bar{\mu}_z$ 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 observed data for each factor. Our prior distribution over the risk neutral parameters $\tilde{\delta}_i$ and $\tilde{\omega}_0^i$ are also normal distributions. For the short rate parameters $\tilde{\delta}_0 \sim \text{N}\left(\mu_\delta, \sigma_\delta^2\right)$, we set the mean of the normal distribution equal to $\mu_\delta = 0.01$ and the variance σ_δ^2 large enough to cover the support of the observed short rate. For the depreciation rate parameters $\tilde{\omega}_0^i \sim \text{N}\left(\mu_\omega, \sigma_\omega^2\right)$, we set the mean of the normal distribution equal to zero $\mu_\omega = 0.005$ and the variance σ_ω^2 large enough to cover the support of the observed depreciation rates.
- For the UFX model, we parameterize the matrix $\tilde{\Phi}_x^*$ as a matrix of eigenvalues.
  - $\tilde{\Phi}_x^*$ is a diagonal matrix of real, ordered eigenvalues. Let $a_1 = -1$ and $b = 1$. We parameterize them as $\tilde{\Phi}_{x,11}^* = a_1 + (b - a_1)U_1$ and $\tilde{\Phi}_{x,jj}^* = a_{j-1} + (b - a_{j-1})U_j$ for $j = 2, \ldots, N$. This transformation ensures that they are increasing and contained in the interval $[-1, 1]$. We then place priors on $\tilde{\Phi}_{x,jj}^*$ via $U_j \sim \text{Beta}\,(12, 12)$.

  For the SFX model, we impose the same prior.
- Consider the UFX model. For each of the free parameters in the factor loadings $\tilde{\delta}_{i,x}$ and/or $\tilde{\delta}_{s,x}$, we place an independent normal prior distribution on each separate element. The prior has mean zero and variance 1.5. The priors are the same for the SFX model.

Appendix C.6 Log-likelihood function

The log-likelihood function is

$$\mathcal{L} \;=\; \log p\left(Y_1,\ldots,Y_T|\theta\right) \;=\; \sum_{t=1}^{T}\log p\left(z_t|z_{t-1},\theta\right)+\sum_{t=1}^{T}\log p\left(\mathrm{Y}_t^{(2)}|z_t;\theta\right)$$

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

$$\mathrm{Y}_t^{(2)} \;=\; C^{(2)}+D^{(2)}z_t+\Sigma_y e_t, \qquad e_t\sim \mathrm{N}\left(0,\mathrm{I}\right),$$

where $C^{(2)}=W_2C$ and $D^{(2)}=W_2D$ 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 C.7 Estimation

Let θ denote all the parameters of the model and define $z_{1:T}=(z_1,\ldots,z_T)$ and $Y_{1:T}=(Y_1,\ldots,Y_T)$. In practice, some data points are missing which implies that some of the factors z_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

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

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

Appendix C.7.1 MCMC algorithm

We provide a brief description of the MCMC algorithm for the UFX model. The MCMC algorithm for the SFX model has similar steps. Let $S_y=\Sigma_y\Sigma_y'$ and $S_z=\Sigma_z\Sigma_z'$ 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 C.7.2. Draw the missing data and location parameters $\left(Y^m_{1:T}, \bar{\mu}_z, \tilde{\delta}_0, \tilde{\omega}^{€}_0, \tilde{\omega}^{£}_0, \tilde{\omega}^{A\$}_0, \tilde{\omega}^{¥}_0\right)$ 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 $z_t = W_1 Y^{o,m}_t$.
- Let $\bar{z}_t = z_t - \bar{\mu}_z$ denote the demeaned factors. We draw the free elements of Φ_z from their full conditional distribution using standard results for Bayesian multiple regression. We write the VAR as a regression model
$$\bar{z}_t = X_t \phi_z + \Sigma_z \varepsilon_t$$
where $\phi_z = \text{vec}(\Phi_z)$ and the regressors X_t contain lagged values of $\bar{z}_{t-1}$. Draws from this model are standard.
- Draw the free elements of S_z from their full conditional using a random-walk Metropolis algorithm. In this step, we avoid conditioning on the parameters S_y, Φ_z by analytically integrating these parameters out of the likelihood.
- Draw the eigenvalues of the risk neutral matrix Φ^*_z from their full conditional using random-walk Metropolis. To avoid conditioning on S_y, Φ_z, we draw from the marginal distribution that analytically integrates these values out of the likelihood.
- Draw the free elements in the factor loadings $\tilde{\delta}_z$ and $\left\{\tilde{\omega}^{€}_z, \tilde{\omega}^{£}_z, \tilde{\omega}^{A\$}_z, \tilde{\omega}^{¥}_z\right\}$ from their full conditional using random-walk Metropolis. To avoid conditioning on S_y, Φ_z, 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 $\bar{\Omega} = \underline{\Omega} + \sum_{t=1}^T e_t e_t^\top$.

Appendix C.7.2 State space form

In our data set, some of the yields contain missing values. We impute them using the Kalman filter. Recall that $C^{(2)} = W_2 C$ and $D^{(2)} = W_2 D$, where $C = \left(\omega_0^\top \; A^\top \; A^{*\top}\right)^\top$ and D collects all the factor loadings. Given that $z_t = Y^{(1)}_t$, we can write the model in VAR form as

$$\begin{pmatrix} Y^{(1)}_t \\ Y^{(2)}_t \end{pmatrix} = \begin{pmatrix} \mu_z \\ C^{(2)} + D^{(2)}\mu_z \end{pmatrix} + \begin{pmatrix} \Phi_z & 0 \\ D^{(2)}\Phi_z & 0 \end{pmatrix} \begin{pmatrix} Y^{(1)}_{t-1} \\ Y^{(2)}_{t-1} \end{pmatrix} + \begin{pmatrix} \Sigma_z & 0 \\ D^{(2)}\Sigma_z & \Sigma_y \end{pmatrix} \begin{pmatrix} w_t \\ e_t \end{pmatrix}$$

Next we translate this system back into the original observed data 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

$$
\begin{aligned}
Y_t &= \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \mu_z \\ C^{(2)} + D^{(2)}\mu_z \end{pmatrix} + \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \Phi_z & 0 \\ D^{(2)}\Phi_z & 0 \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} Y_{t-1} \\
&\quad + \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \Sigma_z & 0 \\ D^{(2)}\Sigma_z & \Sigma_y \end{pmatrix} \begin{pmatrix} w_t \\ e_t \end{pmatrix}
\end{aligned}
$$

This structure implies that Y_t as defined in Appendix C.1 follows a reduced-rank VAR of the form

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

where

$$
\mu_Y = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \mu_z \\ C^{(2)} + D^{(2)}\mu_z \end{pmatrix} \qquad \Phi_Y = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \Phi_z & 0 \\ D^{(2)}\Phi_z & 0 \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}
$$

$$
\Sigma_Y = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}^{-1} \begin{pmatrix} \Sigma_z & 0 \\ D^{(2)}\Sigma_z & \Sigma_y \end{pmatrix} \qquad \varepsilon_{Y,t} = \begin{pmatrix} w_t \\ e_t \end{pmatrix}
$$

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

$$
Y_t = Z\alpha_t + d + u_t \qquad u_t \sim \mathrm{N}(0, H), \tag{C.10}
$$

$$
\alpha_{t+1} = T\alpha_t + c + Rv_t \qquad v_t \sim \mathrm{N}(0, Q). \tag{C.11}
$$

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

Let $\bar{\mu} = \left(\bar{\mu}_x^\top, \tilde{\delta}_0, \tilde{\omega}_0^{€}, \tilde{\omega}_0^{£}, \tilde{\omega}_0^{A\$}, \tilde{\omega}_0^{¥}\right)^\top$ denote the vector of location parameters. The vector of intercepts μ_Y can be written as a linear function of these location parameters

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

We draw the location parameters jointly with the missing data by including them in the state vector. We define the system matrices from (C.10)-(C.11) 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 location parameters 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 parameters jointly.

Appendix D Estimated parameters in the affine models

Table Appendix D.1: Parameter estimates, state dynamics; SFX model.

x_t^S		x_{1t}	x_{2t}	x_{3t}	x_{4t}	x_{5t}	x_{6t}	x_{7t}	x_{8t}	x_{9t}	x_{10t}
	$\bar{\mu}_x \times 1200$		Φ_x								
$x_{1,t+1}$	12.911	0.991	-0.010	0	0	0	0	0	0	0	0
	(0.265)	(0.005)	(0.043)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{2,t+1}$	1.849	-9.41e-04	0.959	0	0	0	0	0	0	0	0
	(0.151)	(0.002)	(0.014)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{3,t+1}$	1.757	-0.010	0.043	0.990	4.70e-04	0	0	0	0	0	0
	(0.246)	(0.004)	(0.035)	(0.007)	(0.023)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{4,t+1}$	0.432	0.002	0.012	-0.027	0.915	0	0	0	0	0	0
	(0.154)	(0.002)	(0.018)	(0.004)	(0.015)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{5,t+1}$	-2.818	-0.013	-0.020	0	0	0.966	0.023	0	0	0	0
	(0.245)	(0.006)	(0.038)	(—)	(—)	(0.012)	(0.037)	(—)	(—)	(—)	(—)
$x_{6,t+1}$	0.381	-0.007	-0.024	0	0	-0.035	0.865	0	0	0	0
	(0.124)	(0.002)	(0.017)	(—)	(—)	(0.006)	(0.018)	(—)	(—)	(—)	(—)
$x_{7,t+1}$	-5.834	-0.015	-0.059	0	0	0	0	0.968	0.014	0	0
	(0.254)	(0.008)	(0.048)	(—)	(—)	(—)	(—)	(0.014)	(0.056)	(—)	(—)
$x_{8,t+1}$	0.401	-0.019	0.017	0	0	0	0	-0.044	0.762	0	0
	(0.118)	(0.005)	(0.028)	(—)	(—)	(—)	(—)	(0.008)	(0.034)	(—)	(—)
$x_{9,t+1}$	6.442	-6.27e-04	0.031	0	0	0	0	0	0	0.983	-0.020
	(0.262)	(0.005)	(0.054)	(—)	(—)	(—)	(—)	(—)	(—)	(0.010)	(0.050)
$x_{10,t+1}$	0.571	0.002	0.208	0	0	0	0	0	0	-0.039	0.711
	(0.133)	(0.003)	(0.032)	(—)	(—)	(—)	(—)	(—)	(—)	(0.007)	(0.030)
			$\Sigma_x \times \sqrt{12} \times 100$								
$x_{1,t+1}$		0.295	0	0	0	0	0	0	0	0	0
		(0.010)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{1,t+1}$		0.023	0.091	0	0	0	0	0	0	0	0
		(0.004)	(0.003)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{1,t+1}$		0.161	0.014	0.145	0	0	0	0	0	0	0
		(0.010)	(0.008)	(0.005)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{1,t+1}$		-0.001	0.020	-0.005	0.092	0	0	0	0	0	0
		(0.006)	(0.005)	(0.004)	(0.003)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{1,t+1}$		0.079	0.018	0.125	0	0.196	0	0	0	0	0
		(0.012)	(0.011)	(0.011)	(—)	(0.007)	(—)	(—)	(—)	(—)	(—)
$x_{1,t+1}$		-0.009	0.019	0	0.053	0.001	0.090	0	0	0	0
		(0.006)	(0.005)	(—)	(0.005)	(0.005)	(0.003)	(—)	(—)	(—)	(—)
$x_{1,t+1}$		0.115	7.06e-04	0.060	0	0.065	0	0.272	0	0	0
		(0.018)	(0.015)	(0.013)	(—)	(0.014)	(—)	(0.011)	(—)	(—)	(—)
$x_{1,t+1}$		-0.047	0.023	0	0.050	0	0.024	-0.024	0.150	0	0
		(0.013)	(0.009)	(—)	(0.008)	(—)	(0.008)	(0.008)	(0.005)	(—)	(—)
$x_{1,t+1}$		0.121	0.037	0.072	0	0.048	0	0.022	0	0.172	0
		(0.012)	(0.010)	(0.009)	(—)	(0.009)	(—)	(0.009)	(—)	(0.006)	(—)
$x_{1,t+1}$		0.010	0.049	0	0.059	0	0.021	0	0.008	-0.015	0.097
		(0.007)	(0.006)	(—)	(0.005)	(—)	(0.005)	(—)	(0.005)	(0.005)	(0.003)

Posterior mean and standard deviation (in parenthesis) of $\bar{\mu}_x, \Phi_x, \Sigma_x$ from the SFX model ($\bar{\mu}_x$ denotes the unconditional mean of the state). The state variables are: $x_{1t} = pc_t^{1,\$}$, $x_{2t} = pc_t^{2,\$}$, $x_{3t} = \Delta_c pc_t^{1,\$} - pc_t^{1,€}$, $x_{4t} = \Delta_c pc_t^{2,€}$, $x_{5t} = \Delta_c pc_t^{1,£}$, $x_{6t} = \Delta_c pc_t^{2,£}$, $x_{7t} = \Delta_c pc_t^{1,A\$}$, $x_{8t} = \Delta_c pc_t^{2,A\$}$, $x_{9t} = \Delta_c pc_t^{1,¥}$, $x_{10t} = \Delta_c pc_t^{2,¥}$.

Table Appendix D.2: Parameter estimates; SFX model.

x_t^S		x_{1t}	x_{2t}	x_{3t}	x_{4t}	x_{5t}	x_{6t}	x_{7t}	x_{8t}	x_{9t}	x_{10t}
	$\bar{\mu}_s \times 1200$		Φ_{sx}								
$\Delta s_{t+1}^{€}$	-8.820	0.019	-0.079	0.319	-0.798	0.018	0.089	-0.011	0.027	-0.009	0.098
	(4.453)	(0.001)	(0.010)	(0.003)	(0.014)	(0.003)	(0.010)	(0.002)	(0.012)	(0.003)	(0.011)
$\Delta s_{t+1}^{£}$	-26.388	-0.042	-0.027	-0.020	-0.068	0.350	-0.765	0.022	0.047	0.039	0.134
	(2.313)	(0.002)	(0.011)	(0.004)	(0.010)	(0.004)	(0.009)	(0.003)	(0.010)	(0.002)	(0.011)
$\Delta s_{t+1}^{A\$}$	-34.904	-0.008	-0.012	-0.048	-0.096	0.039	0.072	0.341	-0.818	0.007	0.130
	(6.777)	(0.003)	(0.014)	(0.004)	(0.012)	(0.003)	(0.008)	(0.003)	(0.012)	(0.003)	(0.013)
$\Delta s_{t+1}^{¥}$	-10.306	0.018	-0.107	-0.005	-0.122	0.007	0.055	-0.002	0.080	0.330	-0.545
	(2.868)	(0.002)	(0.015)	(0.004)	(0.015)	(0.005)	(0.016)	(0.003)	(0.016)	(0.003)	(0.019)
			$\Sigma_{sx} \times \sqrt{12} \times 100$								
$\Delta s_{t+1}^{€}$		11.692	-29.032	-18.544	-16.194	0	0	0	0	0	0
		(5.783)	(5.657)	(11.304)	(9.982)	(—)	(—)	(—)	(—)	(—)	(—)
$\Delta s_{t+1}^{£}$		-29.891	7.691	0	0	54.766	-27.499	0	0	0	0
		(1.904)	(6.937)	(—)	(—)	(4.531)	(12.636)	(—)	(—)	(—)	(—)
$\Delta s_{t+1}^{A\$}$		-22.907	-52.833	0	0	0	0	-47.338	25.051	0	0
		(8.764)	(9.531)	(—)	(—)	(—)	(—)	(7.252)	(9.113)	(—)	(—)
$\Delta s_{t+1}^{¥}$		-12.310	11.530	0	0	0	0	0	0	42.301	-15.294
		(4.391)	(7.387)	(—)	(—)	(—)	(—)	(—)	(—)	(5.594)	(8.009)

Posterior mean and standard deviation (in parenthesis) of the parameters $\bar{\mu}_s, \Phi_{sx}$ and Σ_{sx} of the SFX model. The state variables are: $x_{1t} = pc_t^{1,\$}$, $x_{2t} = pc_t^{2,\$}$, $x_{3t} = \Delta_c pc_t^{1,\$} - pc_t^{1,€}$, $x_{4t} = \Delta_c pc_t^{2,€}$, $x_{5t} = \Delta_c pc_t^{1,£}$, $x_{6t} = \Delta_c pc_t^{2,£}$, $x_{7t} = \Delta_c pc_t^{1,A\$}$, $x_{8t} = \Delta_c pc_t^{2,A\$}$, $x_{9t} = \Delta_c pc_t^{1,¥}$, $x_{10t} = \Delta_c pc_t^{2,¥}$.

Table Appendix D.3: Parameter estimates; SFX model.

$\tilde{x}_t^S$		$\tilde{x}_{1t}$	$\tilde{x}_{2t}$	$\tilde{x}_{3t}$	$\tilde{x}_{4t}$	$\tilde{x}_{5t}$	$\tilde{x}_{6t}$	$\tilde{x}_{7t}$	$\tilde{x}_{8t}$	$\tilde{x}_{9t}$	$\tilde{x}_{10t}$	$\tilde{x}_{11t}$	$\tilde{x}_{12t}$	$\tilde{x}_{13t}$	$\tilde{x}_{14t}$
	$\tilde{\delta}_0$		$\tilde{\delta}_z^{\top}$												
	0.015	1	1	0.819	0.766	0.096	-0.587	-1.876	-0.221	0.531	0.433	0	0	0	0
	(0.003)	(—)	(—)	(0.049)	(0.050)	(0.030)	(0.125)	(0.357)	(0.068)	(0.098)	(0.085)	(—)	(—)	(—)	(—)
	$\tilde{\omega}_0^{€}$		$\tilde{\omega}_z^{€,\top}$												
	0.001	0.402	0.656	1	1	0.252	0.024	1.627	1.307	2.015	0.453	1	0	0	0
	(0.002)	(0.052)	(0.107)	(—)	(—)	(0.021)	(0.043)	(0.374)	(0.145)	(0.178)	(0.184)	(—)	(—)	(—)	(—)
	$\tilde{\omega}_0^{£}$		$\tilde{\omega}_z^{£,\top}$												
	-0.020	1.592	-1.200	0.836	1.887	1	1	-1.545	-0.879	-1.744	-0.196	0	1	0	0
	(0.004)	(0.118)	(0.261)	(0.137)	(0.132)	(—)	(—)	(0.331)	(0.095)	(0.184)	(0.180)	(—)	(—)	(—)	(—)
	$\tilde{\omega}_0^{A\$}$		$\tilde{\omega}_z^{A\$,\prime}$												
	-0.025	1.764	-2.797	-1.135	-1.356	-0.527	-0.734	1	1	1.732	-3.121	0	0	1	0
	(0.010)	(0.168)	(0.235)	(0.088)	(0.080)	(0.032)	(0.082)	(—)	(—)	(0.110)	(0.476)	(—)	(—)	(—)	(—)
	$\tilde{\omega}_0^{¥}$		$\tilde{\omega}_z^{¥,\prime}$												
	-4.28e-05	0.959	0.923	1.847	2.368	0.544	-0.350	-0.679	0.366	1	1	0	0	0	1
	(0.005)	(0.092)	(0.108)	(0.082)	(0.058)	(0.027)	(0.083)	(0.229)	(0.051)	(—)	(—)	(—)	(—)	(—)	(—)
			$\tilde{\Phi}_x^*$												
$\tilde{x}_{1,t+1}$		0.999	0	0	0	0	0	0	0	0	0	0	0	0	0
		(1.70e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{2,t+1}$		0	0.996	0	0	0	0	0	0	0	0	0	0	0	0
		(—)	(2.23e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{3,t+1}$		0	0	0.992	0	0	0	0	0	0	0	0	0	0	0
		(—)	(—)	(4.90e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{4,t+1}$		0	0	0	0.988	0	0	0	0	0	0	0	0	0	0
		(—)	(—)	(—)	(5.35e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{5,t+1}$		0	0	0	0	0.982	0	0	0	0	0	0	0	0	0
		(—)	(—)	(—)	(—)	(4.39e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{6,t+1}$		0	0	0	0	0	0.974	0	0	0	0	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(7.38e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{7,t+1}$		0	0	0	0	0	0	0.963	0	0	0	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(—)	(0.001)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{8,t+1}$		0	0	0	0	0	0	0	0.957	0	0	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(—)	(—)	(0.002)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{9,t+1}$		0	0	0	0	0	0	0	0	0.943	0	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(0.002)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{10,t+1}$		0	0	0	0	0	0	0	0	0	0.870	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(0.006)	(—)	(—)	(—)	(—)

Posterior mean and standard deviation (in parenthesis) of the parameters of the SFX model. The state variables are reported under the latent factor rotation, $\tilde{x}_t$, which corresponds to the diagonal persistence matrix $\tilde{\Phi}_x^*$. Parameterization is discussed in Appendix C.4.1.

Table Appendix D.4: Parameter estimates; UFX model.

x_t^U		x_{1t}	x_{2t}	x_{3t}	x_{4t}	x_{5t}	x_{6t}	x_{7t}	x_{8t}	x_{9t}	x_{10t}
	$\bar{\mu}_x \times 1200$		Φ_x								
$x_{1,t+1}$	11.836	0.991	-0.010	0	0	0	0	0	0	0	0
	(0.119)	(0.005)	(0.045)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{2,t+1}$	1.866	-9.71e-04	0.958	0	0	0	0	0	0	0	0
	(0.088)	(0.002)	(0.014)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{3,t+1}$	1.480	-0.010	0.044	0.990	0.001	0	0	0	0	0	0
	(0.117)	(0.004)	(0.036)	(0.007)	(0.023)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{4,t+1}$	0.462	0.003	0.014	-0.028	0.912	0	0	0	0	0	0
	(0.096)	(0.002)	(0.018)	(0.004)	(0.015)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{5,t+1}$	-2.514	-0.013	-0.022	0	0	0.965	0.025	0	0	0	0
	(0.117)	(0.005)	(0.039)	(—)	(—)	(0.012)	(0.038)	(—)	(—)	(—)	(—)
$x_{6,t+1}$	0.330	-0.008	-0.022	0	0	-0.037	0.859	0	0	0	0
	(0.084)	(0.002)	(0.017)	(—)	(—)	(0.006)	(0.018)	(—)	(—)	(—)	(—)
$x_{7,t+1}$	-5.691	-0.013	-0.065	0	0	0	0	0.972	0.040	0	0
	(0.118)	(0.008)	(0.049)	(—)	(—)	(—)	(—)	(0.014)	(0.055)	(—)	(—)
$x_{8,t+1}$	0.499	-0.018	0.028	0	0	0	0	-0.045	0.742	0	0
	(0.080)	(0.005)	(0.027)	(—)	(—)	(—)	(—)	(0.008)	(0.030)	(—)	(—)
$x_{9,t+1}$	6.997	-1.89e-04	0.026	0	0	0	0	0	0	0.981	-0.018
	(0.118)	(0.005)	(0.054)	(—)	(—)	(—)	(—)	(—)	(—)	(0.011)	(0.049)
$x_{10,t+1}$	0.489	0.003	0.214	0	0	0	0	0	0	-0.041	0.705
	(0.077)	(0.003)	(0.032)	(—)	(—)	(—)	(—)	(—)	(—)	(0.007)	(0.030)
			$\Sigma_x \times \sqrt{12} \times 100$								
$x_{1,t+1}$		0.305	0	0	0	0	0	0	0	0	0
		(0.010)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{2,t+1}$		0.023	0.091	0	0	0	0	0	0	0	0
		(0.005)	(0.003)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{3,t+1}$		0.159	0.011	0.148	0	0	0	0	0	0	0
		(0.009)	(0.007)	(0.005)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{4,t+1}$		-0.006	0.020	-0.004	0.094	0	0	0	0	0	0
		(0.006)	(0.005)	(0.004)	(0.003)	(—)	(—)	(—)	(—)	(—)	(—)
$x_{5,t+1}$		0.063	0.013	0.133	0	0.201	0	0	0	0	0
		(0.013)	(0.012)	(0.012)	(—)	(0.007)	(—)	(—)	(—)	(—)	(—)
$x_{6,t+1}$		-0.016	0.019	0	0.055	0.003	0.090	0	0	0	0
		(0.006)	(0.005)	(—)	(0.005)	(0.004)	(0.003)	(—)	(—)	(—)	(—)
$x_{7,t+1}$		0.106	0.003	0.069	0	0.070	0	0.271	0	0	0
		(0.017)	(0.014)	(0.014)	(—)	(0.013)	(—)	(0.009)	(—)	(—)	(—)
$x_{8,t+1}$		-0.044	0.025	0	0.051	0	0.027	-0.024	0.149	0	0
		(0.012)	(0.008)	(—)	(0.008)	(—)	(0.008)	(0.008)	(0.005)	(—)	(—)
$x_{9,t+1}$		0.111	0.032	0.080	0	0.054	0	0.021	0	0.173	0
		(0.011)	(0.010)	(0.010)	(—)	(0.009)	(—)	(0.008)	(—)	(0.006)	(—)
$x_{10,t+1}$		0.004	0.049	0	0.060	0	0.023	0	0.010	-0.014	0.097
		(0.007)	(0.006)	(—)	(0.006)	(—)	(0.005)	(—)	(0.005)	(0.005)	(0.003)

Posterior mean and standard deviation (in parenthesis) of the parameters $\bar{\mu}_x, \Phi_x$, and Σ_x of the UFX model. The state variables are: $x_{1t} = pc_t^{1,\$}$, $x_{2t} = pc_t^{2,\$}$, $x_{3t} = \Delta_c pc_t^{1,\$} - pc_t^{1,€}$, $x_{4t} = \Delta_c pc_t^{2,€}$, $x_{5t} = \Delta_c pc_t^{1,£}$, $x_{6t} = \Delta_c pc_t^{2,£}$, $x_{7t} = \Delta_c pc_t^{1,A\$}$, $x_{8t} = \Delta_c pc_t^{2,A\$}$, $x_{9t} = \Delta_c pc_t^{1,¥}$, $x_{10t} = \Delta_c pc_t^{2,¥}$.

Table Appendix D.5: Parameter estimates; UFX model.

x_t^U		x_{1t}	x_{2t}	x_{3t}	x_{4t}	x_{5t}	x_{6t}	x_{7t}	x_{8t}	x_{9t}	x_{10t}
	$\bar{\mu}_s \times 1200$		Φ_{sx}								
$\Delta s^{€}_{t+1}$	1.288	0.189	0.336	-0.403	0.473	0	0	0	0	0	0
	(1.803)	(0.231)	(1.355)	(0.406)	(1.174)	(—)	(—)	(—)	(—)	(—)	(—)
$\Delta s^{£}_{t+1}$	-0.742	0.064	0.415	0	0	-0.076	-0.380	0	0	0	0
	(1.720)	(0.224)	(1.226)	(—)	(—)	(0.449)	(1.218)	(—)	(—)	(—)	(—)
$\Delta s^{A\$}_{t+1}$	-0.533	-0.671	1.786	0	0	0	0	-0.994	-3.126	0	0
	(1.949)	(0.273)	(1.382)	(—)	(—)	(—)	(—)	(0.435)	(1.458)	(—)	(—)
$\Delta s^{¥}_{t+1}$	1.315	0.330	8.564	0	0	0	0	0	0	-0.796	-4.751
	(1.906)	(0.261)	(1.581)	(—)	(—)	(—)	(—)	(—)	(—)	(0.546)	(1.543)
			$\Sigma_{sx} \times \sqrt{12} \times 100$								
$\Delta s^{€}_{t+1}$		-1.213	0.262	-1.173	-0.278	0	0	0	0	0	0
		(0.528)	(0.508)	(0.365)	(0.377)	(—)	(—)	(—)	(—)	(—)	(—)
$\Delta s^{£}_{t+1}$		-1.361	0.614	0	0	0.389	0.498	0	0	0	0
		(0.501)	(0.469)	(—)	(—)	(0.373)	(0.377)	(—)	(—)	(—)	(—)
$\Delta s^{A\$}_{t+1}$		0.318	-0.933	0	0	0	0	0.399	0.087	0	0
		(0.535)	(0.519)	(—)	(—)	(—)	(—)	(0.506)	(0.534)	(—)	(—)
$\Delta s^{¥}_{t+1}$		-1.899	0.591	0	0	0	0	0	0	-1.097	-0.637
		(0.519)	(0.514)	(—)	(—)	(—)	(—)	(—)	(—)	(0.470)	(0.463)
			$\Sigma_s \times \sqrt{12} \times 100$								
		$\eta^{€}_t$	$\eta^{£}_t$	$\eta^{A\$}_t$	$\eta^{¥}_t$						
$\Delta s^{€}_{t+1}$		10.516	0	0	0						
		(0.369)	(—)	(—)	(—)						
$\Delta s^{£}_{t+1}$		6.671	7.510	0	0						
		(0.421)	(0.250)	(—)	(—)						
$\Delta s^{A\$}_{t+1}$		4.558	1.573	10.317	0						
		(0.549)	(0.497)	(0.366)	(—)						
$\Delta s^{¥}_{t+1}$		4.690	0.005	-0.463	9.469						
		(0.485)	(0.462)	(0.469)	(0.329)						

Posterior mean and standard deviation (in parenthesis) of the parameters $\bar{\mu}_s, \Phi_{sx}, \Sigma_{sx}$, and Σ_s in the UFX model. The state variables are: $x_{1t} = pc_t^{1,\$}$, $x_{2t} = pc_t^{2,\$}$, $x_{3t} = \Delta_c pc_t^{1,\$} - pc_t^{1,€}$, $x_{4t} = \Delta_c pc_t^{2,€}$, $x_{5t} = \Delta_c pc_t^{1,£}$, $x_{6t} = \Delta_c pc_t^{2,£}$, $x_{7t} = \Delta_c pc_t^{1,A\$}$, $x_{8t} = \Delta_c pc_t^{2,A\$}$, $x_{9t} = \Delta_c pc_t^{1,¥}$, $x_{10t} = \Delta_c pc_t^{2,¥}$.

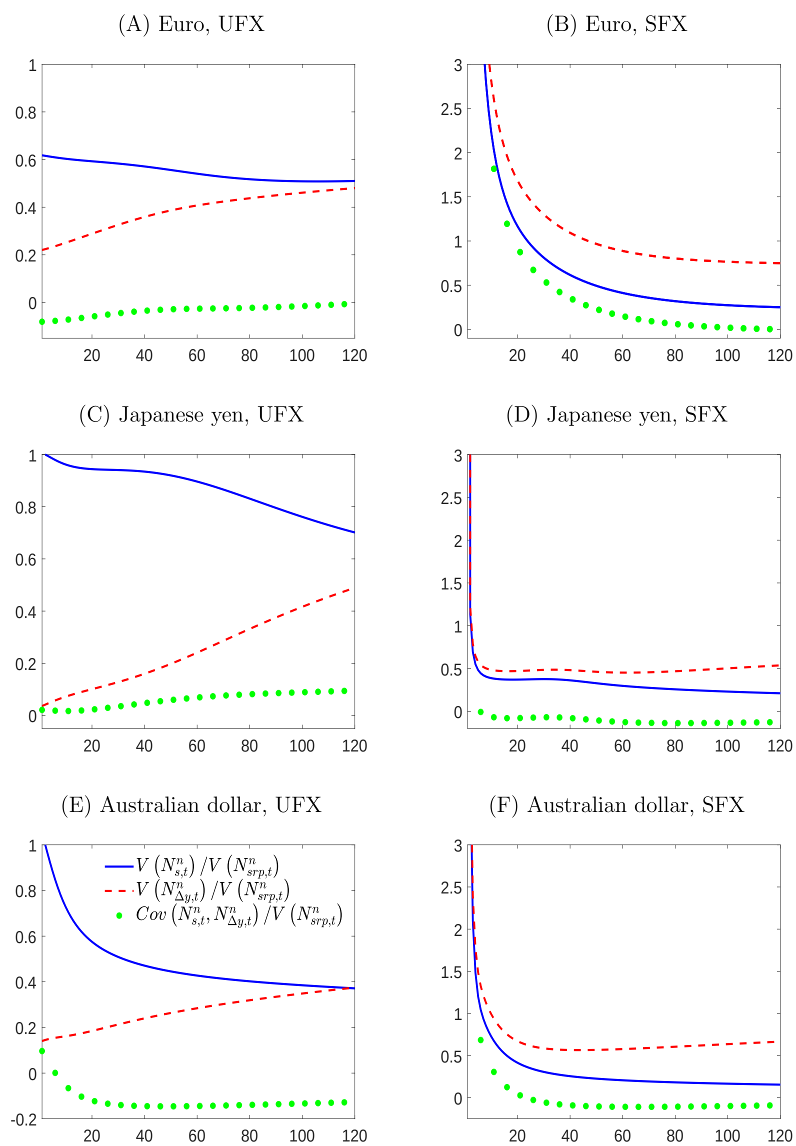
Table Appendix D.6: Parameter estimates; UFX model.

$\tilde{x}_t^W$		$\tilde{x}_{1t}$	$\tilde{x}_{2t}$	$\tilde{x}_{3t}$	$\tilde{x}_{4t}$	$\tilde{x}_{5t}$	$\tilde{x}_{6t}$	$\tilde{x}_{7t}$	$\tilde{x}_{8t}$	$\tilde{x}_{9t}$	$\tilde{x}_{10t}$	$\tilde{x}_{11t}$	$\tilde{x}_{12t}$	$\tilde{x}_{13t}$	$\tilde{x}_{14t}$
	$\tilde{\delta}_0$		$\tilde{\delta}_z^{\top}$												
	0.017	1	1	-0.477	-0.552	-0.284	-0.692	-1.298	0.167	0.655	0.262	0	0	0	0
	(0.002)	(—)	(—)	(0.078)	(0.072)	(0.051)	(0.082)	(0.228)	(0.242)	(0.121)	(0.099)	(—)	(—)	(—)	(—)
	$\tilde{\omega}_0^{€}$		$\tilde{\omega}_z^{€,\top}$												
	0.005	0.283	-0.003	1	1	0.256	0.065	1.428	1.237	1.233	0.942	1	0	0	0
	(6.94e-04)	(0.075)	(0.160)	(—)	(—)	(0.053)	(0.082)	(0.229)	(0.066)	(0.113)	(0.178)	(—)	(—)	(—)	(—)
	$\tilde{\omega}_0^{£}$		$\tilde{\omega}_z^{£,\top}$												
	0.009	1.146	2.003	-0.223	1.225	1	1	-1.702	-1.020	-1.337	-0.098	0	1	0	0
	(0.002)	(0.092)	(0.109)	(0.206)	(0.222)	(—)	(—)	(0.335)	(0.193)	(0.203)	(0.239)	(—)	(—)	(—)	(—)
	$\tilde{\omega}_0^{A\$}$		$\tilde{\omega}_z^{A\$,\prime}$												
	0.003	0.827	2.100	-2.064	-2.136	-0.563	-0.541	1	1	1.040	-3.093	0	0	1	0
	(0.001)	(0.029)	(0.130)	(0.297)	(0.445)	(0.116)	(0.127)	(—)	(—)	(0.140)	(0.464)	(—)	(—)	(—)	(—)
	$\tilde{\omega}_0^{¥}$		$\tilde{\omega}_z^{¥,\prime}$												
	0.010	0.746	0.472	1.507	2.233	0.316	-0.576	-0.417	0.705	1	1	0	0	0	1
	(0.002)	(0.052)	(0.098)	(0.139)	(0.217)	(0.098)	(0.153)	(0.320)	(0.188)	(—)	(—)	(—)	(—)	(—)	(—)
			$\tilde{\Phi}_x^*$												
$\tilde{x}_{1,t+1}$		0.998	0	0	0	0	0	0	0	0	0	0	0	0	0
		(2.37e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{2,t+1}$		0	0.996	0	0	0	0	0	0	0	0	0	0	0	0
		(—)	(2.03e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{3,t+1}$		0	0	0.992	0	0	0	0	0	0	0	0	0	0	0
		(—)	(—)	(4.72e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{4,t+1}$		0	0	0	0.988	0	0	0	0	0	0	0	0	0	0
		(—)	(—)	(—)	(6.46e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{5,t+1}$		0	0	0	0	0.981	0	0	0	0	0	0	0	0	0
		(—)	(—)	(—)	(—)	(7.71e-04)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{6,t+1}$		0	0	0	0	0	0.974	0	0	0	0	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(0.001)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{7,t+1}$		0	0	0	0	0	0	0.964	0	0	0	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(—)	(0.002)	(—)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{8,t+1}$		0	0	0	0	0	0	0	0.958	0	0	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(—)	(—)	(0.001)	(—)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{9,t+1}$		0	0	0	0	0	0	0	0	0.943	0	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(0.003)	(—)	(—)	(—)	(—)	(—)
$\tilde{x}_{10,t+1}$		0	0	0	0	0	0	0	0	0	0.873	0	0	0	0
		(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(—)	(0.007)	(—)	(—)	(—)	(—)

Posterior mean and standard deviation (in parenthesis) of the parameters of the UFX model. The state variables are reported under the latent factor rotation, $\tilde{x}_t$, which corresponds to the diagonal persistence matrix $\tilde{\Phi}_x^*$. Parameterization is discussed in Appendix C.4.1.

Appendix E Additional results on news decomposition

Figure Appendix E.1
News-based decomposition of currency risk premiums



Notes: We plot the 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*. We use the UFX and SFX models to compute the decomposition.*