HONG KONG INSTITUTE FOR MONETARY AND FINANCIAL RESEARCH

WHAT 200 YEARS OF DATA TELL US ABOUT THE PREDICTIVE VARIANCE OF LONG-TERM BONDS?

Pasquale Della Corte, Can Gao, Daniel P.A. Preve, and Giorgio Valente

HKIMR Working Paper No. 01/2025

February 2025

HONG KONG ACADEMY OF FINANCE

金融學院

Hong Kong Institute for Monetary and Financial Research
香港貨幣及金融研究中心
(a company incorporated with limited liability)

What 200 Years of Data Tell Us About the Predictive Variance of Long-Term Bonds?*

Pasquale Della Corte

Imperial College London

& CEPR

Can Gao

University of St.Gallen

& Swiss Finance Institute

Daniel P.A. Preve

SMU

Giorgio Valente

HKIMR

February 2025

*The authors would like to thank Juan Antolin Diaz, Gino Cenedese, Yu-chin Chen, Max Croce, Bernard Dumas, Ľuboš Pástor, Fabio Trojani, conference participants at the 2018 Workshop on Bayesian Methods in Finance in Singapore, 2018 Belgrade Symposium in Economics and Finance, 2021 Asset Pricing Conference by Long-Term Investors@UniTo, 2022 ABFER Annual Conference in Singapore, as well as seminar participants at the Bank of England and Fulcrum Asset Management for their helpful and constructive comments. The views expressed in this paper are those of the authors and are not necessarily those of the Bank of England, the Hong Kong Monetary Authority, the Hong Kong Academy of Finance or the Hong Kong Institute for Monetary and Financial Research (HKIMR). The above-mentioned entities except the authors take no responsibility for inaccuracies or omissions contained in the paper. Pasquale Della Corte acknowledges the gracious hospitality of the HKIMR, where part of this research was conducted.

Pasquale Della Corte is with Imperial College Business School, Imperial College London and the Centre for Economic Policy Research (CEPR), London. Email: p.dellacorte@imperial.ac.uk. Can Gao is with Swiss Finance Institute (SFI) and University of St.Gallen. Email: can.gao@unisg.ch. Daniel P.A. Preve is with the School of Economics, Singapore Management University. Email: dpreve@smu.edu.sg. Giorgio Valente is with Hong Kong Institute for Monetary and Financial Research, Hong Kong Monetary Authority, Hong Kong. Email: gvalente@hkma.gov.hk.

What 200 Years of Data Tell Us About the Predictive Variance of Long-Term Bonds?

Abstract

This paper investigates the long-horizon predictive variance of an international bond strategy where a U.S. investor holds unhedged positions in constant-maturity long-term foreign bonds funded at domestic short-term interest rates. Using over two centuries of data from major economies, the study finds that predictive variance grows with the investment horizon, driven primarily by uncertainties in interest rate differentials and exchange rate returns, which outweigh mean reversion effects. The analysis, incorporating both observable and unobservable predictors, highlights that unobservable predictors linked to shifts in monetary and exchange rate regimes are the dominant source of long-term risk, offering fresh insights into international bond investment strategies.

Keywords: Currency risk, Long-term bonds, Predictability, Long-term investments.
JEL Classification: F31, G12, G15.

1 Introduction

Over the past two decades, there has been a greater appetite for long-term bonds as noted by leading economic newspapers like the *Wall Street Journal* and the *Financial Times* (e.g., Allen, 2017; Sindreu, 2017). On the one hand, with short-term interest rates reaching their historical lows due to unconventional monetary policies, investors in search for higher yields have tilted their allocation towards long-term bonds, often using leveraged positions. On the other hand, with post-crisis reforms aimed at financial stability, institutional investors have increasingly turned to long-term government bonds, which require no additional regulatory capital, are less complex than centrally-cleared derivatives to hedge interest rate risk, and serve as key collateral in short-term borrowing markets. These trends are reflected in the growth and composition of the global bond market. While its outstanding amount has expanded from about $30 trillion in 2000 to approximately $104 trillion in 2019, the ratio of long-term bonds to gross domestic product has nearly doubled from 38% in 2000 to 64% in 2019 (e.g., Bogdanova, Chan, Micic and von Peter, 2021).

Despite their growing importance as a key asset class, the risk profile of long-term bonds remains largely unexplored, especially from an ex-ante perspective. Understanding such risk is even more complex when cross-country differences in monetary policies, as observed in recent years, drive investors to increasingly rely on short-term dollar funding (e.g., CGFS, 2020), thus heightening exposure to both interest rate and foreign exchange fluctuations. An example occurred during the 2022 UK mini-budget crisis, which triggered a sharp rise in bond yields alongside a large depreciation of the pound sterling against the dollar. This episode created significant volatility for international bondholders, resulting from the compounding effect of interest rate and exchange rate fluctuations.

In this paper, we attempt to fill this gap in the literature by assessing the predictive variance of an international bond strategy in which a US investor holds an unhedged position in a constant-maturity foreign bond funded at the domestic short-term interest rate over the long horizon. To preview our results, we find that holding long-term bonds generates substantial predictive variance, which grows significantly over longer horizons. This upward-sloping behaviour is predominantly driven by uncertainties about future exchange rates and interest rate differentials, which outweigh the effects of mean reversion in returns. Also, these results leverage more than 200 years of data and are robust across major economies. Our findings provide an alternative perspective to recent studies on the risk profile of long-term bonds. For example, Meyer, Reinhart and Trebesch (2022) use two centuries of data covering a large cross-section of countries and document that, like the equity market, long-term government bonds have historically generated high excess returns with relatively low volatility. In a similar vein, Viceira and Wang (2018) investigate global portfolio diversification for long-horizon investors and report that long-term bonds can lower risk. While these studies rely on ex-post results, our analysis adopts an ex-ante perspective that incorporates both parameter uncertainty and imperfect predictability akin to Pástor and Stambaugh (2012) and Avramov, Cederburg and Lučivjanská (2018).

Our analysis starts by decomposing the returns of an international bond strategy into three components: foreign bond excess returns, real interest rate differentials, and real exchange rate returns. It then adopts the predictive variance as a notion of risk as this is what really matters to investors for their long-horizon portfolio decisions. In particular, when forming expectations about future returns, investors ignore the true data generating process and, by relying on potentially misspecified empirical models, observable predictors may deliver imperfect forecasts. This implies that an investor's predictive variance differs substantially from the actual variance as the former encompasses a range of uncertainties which are absent in the latter (e.g., Pástor and Stambaugh, 2012; Avramov et al., 2018). These uncertainties are important for long-term investors as they are likely to offset the effects of mean reversion in returns for longer investment horizons, even in the presence of return predictability.[1]

We carry out an empirical investigation exploring the long-term predictive variances of returns from investments in long-term bonds denominated in major currencies over the past

[1]The effect of these uncertainties on the predictive variance may be even stronger if returns are only partially predictable.

two centuries. In the spirit of Pástor and Stambaugh (2009), we estimate these long-term predictive variances in an environment with imperfect predictability, by allowing unobserved predictors to join a set of observable predictors to help forecast the strategy returns at different horizons. As the variables to be forecast in our setting relate to interest rates and exchange rates, we conjecture that the unobserved predictors in our framework can be potentially associated with changes in monetary and exchange rate regimes that have occurred over the past two centuries and are not already captured by the set of observable predictors. Furthermore, we derive in closed form a range of uncertainties that affect the predictive variance for long horizons, in addition to the mean reversion component due to the predictability of returns, building upon and extending the theoretical frameworks proposed by Pástor and Stambaugh (2012) and Avramov, Cederburg and Lučivjanská (2018).

The estimations lead to a host of novel results. First, over the full sample period and across all countries, the predictive variance of the bond investment strategy is found to be increasing with the investment horizon and this is mainly due to a growing predictive variance for both short-term interest rate differentials and real exchange rate returns. Overall, the predictive variance of real exchange rate returns exhibits the largest long-horizon value, followed by that of interest rate differentials, while the predictive variance of bond excess returns in foreign currency does not vary much across investment horizons. The predictive covariances are all negative, especially between interest rate differentials and real exchange rate returns, which is consistently negative across currency pairs and decreasing over the investment horizon. This finding suggests that the predictive co-movement between foreign bond excess returns and real interest rate differential is less important in determining the long-term risk profile of the strategy. The results also suggest that the predictive co-movements between interest rate differentials and real exchange rate returns are important in the long-run as they tend to reduce the overall expected risk of the strategy, especially at longer horizons.

Second, after decomposing the predictive variance of the strategy returns into its key constituents, we observe that in all cases the uncertainty about future returns unambiguously plays the leading role. All other components, especially the one associated with mean reversion due to return predictability, are negligible in size and do not provide any improvement in the risk profile of the strategy at longer horizons. Put differently, when foreign bond excess returns, real interest rate differentials, and real exchange rate returns predictability are taken seriously into account, the notion that bond returns are less volatile in the long-run does not apply. The range of uncertainties that affect the predictive variance more than offset any potential benefit originating from mean reversion in returns. Upon further exploration, across all of the currency pairs investigated, the uncertainty about future returns is mainly due to the component of the predictive variance pertaining to the unobserved predictors. The uncertainty associated with the expected future values of the observable predictors is non-negligible but substantially smaller than the one documented for unobserved predictors. Furthermore, as the uncertainty about future returns originating from interest rate differentials and exchange rate returns is fairly similar, the shape of the predictive variance over longer horizons can be interpreted as spurring from changes in monetary and exchange rate regimes that are not captured by the set of observable predictors. Also the negative impact arising from mean reversion is not large enough to offset future uncertainty.

Third, for our baseline exercise involving the dollar relative to the pound sterling, we also estimate the predictive variance using an expanding window that starts with 100 years of data and then progressively incorporates additional years until the end of the sample. The analysis reveals substantial variation in the slope of the predictive variance. It was particularly steep around the Great War and later during the Black Wednesday. Also, while the slope of predictive variance of interest rate differentials has been declining over time, the slope of predictive variance of real exchange rate returns has been fairly stable over time. Interestingly, the predictive variance of foreign bond excess returns has been particularly steep during the second World War and again when the UK was sick man of Europe. Finally, a number of robustness exercises using different priors, removing the observable predictors, working with the recent floating period, confirms our main findings.

Our study is related to various strands of the literature. First, it speaks to a vast body of research that studies bond returns and yields over various investment horizons. Several studies find evidence of predictability for US bond yields and returns provided by interest forward rates, macroeconomic fundamentals and principal components of bond yields (e.g., Fama and Bliss, 1987; Cochrane and Piazzesi, 2005; Della Corte, Sarno and Thornton, 2008a; Ludvigson and Ng, 2009; Joslin, Priebsch and Singleton, 2014). Evidence of predictability for non-US bond returns is overall less pervasive than for US bond returns (Ilmanen, 1995). More recently, Meyer et al. (2022) explores the performance of external sovereign bonds over two centuries, analyzing cycles of boom and bust. Their findings reveal that sovereign bonds have offered sufficiently high returns to compensate for risks. As we are concerned with the risk of investing in foreign long-term bonds, the predictability of FX returns is also important. This literature, however, has not reached a consensus as to whether FX returns are predictable (Meese and Rogoff, 1983; Engel and West, 2005). In our study, we allow for predictability in all components of the strategy returns but we also take into account the possibility that such predictability is imperfect akin to Pástor and Stambaugh (2009).

The second body of research we build on relates to the measurement of risk associated with portfolio investments over long horizons, and the implications for asset allocation. In addition to the pioneering works by Samuelson (1969) and Merton (1969), who show that investors should choose the same asset allocation regardless of investment horizon whenever asset returns are unpredictable, our empirical analysis builds upon the results of Siegel (1992, 2008), Barberis (2000) and Campbell and Viceira (2002, 2005). These studies show that, in the presence of return predictability, the perceived variability of asset returns is lower for longer horizons because of the effect mean reversion of expected returns has on the long-horizon variance. The two studies that are closest to our empirical investigation are the ones by Pástor and Stambaugh (2012) and Avramov et al. (2018), who show that asset returns are more volatile over longer horizons if the predictive variance of returns is used as the main notion of long-horizon risk. This novel result is due to the presence of an assortment of uncertainties that are explicitly included in the framework for the predictive variance, but not for the true variance. We improve upon these works in several important ways: We first show how the long-horizon predictive covariances associated with a predictive system that includes both observable and unobserved predictors can be decomposed into five main components, with accompanying closed-form expressions. Thus, in contrast to Pástor and Stambaugh (2012) and Avramov et al. (2018), who focus on the long-horizon predictive variance of a single asset, our focus is on obtaining an informative decomposition for the long-horizon predictive covariance of multiple assets. We then apply this framework to a multiple asset case, allowing for imperfect predictability. This permits us to gain further insight into the long-horizon predictive variance of the strategy returns and to directly link the main sources of uncertainty, including mean reversion in expected returns, to the estimated parameters of the predictive system. Recent research by Froot (2019) investigates the effectiveness of currency hedging for international investments over different time horizons. The study finds that currency hedging reduces portfolio variance at short horizons but may increase portfolio variance at long horizons. This is due to the greater influence of unexpected changes in relative inflation and interest rates between countries over long periods, which can introduce additional volatility and make hedging less beneficial or even counterproductive for long-term investors.

Our study is structured as follows. Section 2 presents our framework for deriving and computing long-horizon predictive covariances in the presence of imperfect predictability, and discusses the theoretical findings. Section 3 shows the components of the strategy returns and introduces their long-horizon predictive variance. It also describes the long-span data used in the empirical investigation and reports some preliminary statistics. Section 4 reports the main results, Section 5 discusses a number of robustness checks, whereas Section 6 concludes. A separate Internet Appendix describes the Bayesian estimation and presents the derivation of the decomposition for the long-horizon predictive covariance.

2 Framework

This section begins by introducing a simple international bond strategy similar to Andrews, Colacito, Croce and Gavazzoni (2024). It then introduces a framework for assessing its multiperiod predictive variance that extends the methodology outlined in Pástor and Stambaugh

(2012) and Avramov, Cederburg and Lučivjanská (2018). Our focus is to develop a parsimonious predictive system that incorporates both observable and unobservable predictors. We then break down the multiperiod predictive variance into five key components, each obtained through closed-form solutions.

2.1 An International Bond Strategy

Consider a simple strategy in which a US investor holds, in each month t, a constant-maturity long-term bond denominated in foreign currency, funded through borrowing at the domestic short-term interest rate. The investor's excess return between months t and $t+1$ is then given by

$$rx_{t+1} = y^{\star}_{t+1} + e_{t+1} - i_{t+1},$$

where $y^{\star}_{t+1}$ is the one-month return on a constant maturity long-term bond denominated in foreign currency, e_{t+1} is the one-month nominal exchange rate return, and i_{t+1} is the one-month return on a short-term bond denominated in domestic currency. All returns are in logs and defined between months t and $t+1$.[2]

In our empirical analysis, we work with more than two centuries of data, and it is convenient to rewrite the above excess return in terms of real quantities as

$$rx_{t+1} = \underbrace{\underbrace{y^{\star}_{t+1} - i^{\star}_{t+1}}_{\substack{\textit{foreign bond} \\ \textit{excess return}}}}_{r_{1,t+1}} + \underbrace{\underbrace{(i^{\star}_{t+1} - \rho^{\star}_{t+1}) - (i_{t+1} - \rho_{t+1})}_{\substack{\textit{real interest rate differential} \\ \textit{between foreign and domestic}}}}_{r_{2,t+1}} + \underbrace{\underbrace{e_{t+1} + \rho^{\star}_{t+1} - \rho_{t+1}}_{\substack{\textit{real exchange} \\ \textit{rate return}}}}_{r_{3,t+1}}, \tag{1}$$

where $i^{\star}_{t+1}$ denotes the one-month return on a short-term bond in foreign currency, ρ_{t+1} refers to the one-month domestic inflation rate, and $\rho^{\star}_{t+1}$ indicates the one-month foreign inflation rate. This decomposition shows how foreign bond excess returns, ex-post real interest rate

[2] We use i_{t+1} as opposed to i_t to indicate the monthly interest rate between times t and $t+1$ as our empirical analysis uses monthly returns on a Treasury Bill index based on three-month maturity instruments.

differentials, and real exchange rate returns contribute to the overall excess returns, thereby capturing the impact of local economic conditions, monetary policy divergence, and exchange rate behavior. To ease the notation, we denote the foreign bond excess return as $r_{1,t+1}$, the real interest rate differential between foreign and domestic country as $r_{2,t+1}$, and the real exchange rate return as $r_{3,t+1}$.

2.2 Predictive System with Imperfect Predictability

We jointly model the terms on the right side of Equation (1) using a linear predictive system

$$r_{t+1} = a + bx_t + \pi_t + u_{t+1} \tag{2}$$

$$x_{t+1} = \theta + \gamma x_t + v_{t+1} \tag{3}$$

$$\pi_{t+1} = \delta \pi_t + \eta_{t+1}, \tag{4}$$

where r_{t+1} is a vector that stacks together foreign bond excess returns, real interest rate differentials, and real exchange rate returns, x_t and π_t are vectors of observable and unobservable predictors, a and θ are vectors of intercepts, while b, γ and δ are diagonal matrices of slope coefficients. All vectors and matrices have a three-dimensional form. As described later in Section 3, the set of observable predictors includes the term spread for the foreign bond excess return, the real output growth differential for the real interest rate differential, and the nominal interest rate differential for the real exchange rate return. These choices are guided by data availability over long sample periods. Also, to ensure stability in our predictive system, we restrict the diagonal elements of γ and δ so that $-1 < \gamma_i < 1$ and $0 < \delta_i < 1$ for $i = \{1, 2, 3\}$. These constraints ensure that the observable predictors in x_t exhibit a mean reverting behavior, while the unobservable predictors in π_t follow a stationary process and positively contribute to return predictability. Note that the process for π_t is without a drift since we already have the constant a in the return equation.

The vectors of residuals are independent and identically normally distributed as

$$\begin{bmatrix} u_t \\ v_t \\ \eta_t \end{bmatrix} \overset{iid}{\sim} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{uu} & \Sigma'_{vu} & \Sigma'_{\eta u} \\ \Sigma_{vu} & \Sigma_{vv} & \Sigma'_{\eta v} \\ \Sigma_{\eta u} & \Sigma_{\eta v} & \Sigma_{\eta\eta} \end{bmatrix} \right), \tag{5}$$

where Σ_{uu} is the covariance matrix of the unexpected returns in u_t, Σ_{vv} is the covariance matrix of shocks affecting the observable predictors in v_t, and $\Sigma_{\eta\eta}$ is the covariance matrix of shocks contaminating the unobservable predictors in η_t. The off-diagonal matrices, such as Σ_{vu} or $\Sigma_{\eta u}$, are cross-equation covariances matrices capturing the interaction between unexpected returns and shocks to observable or unobservable predictors. All these matrices have a three-dimensional form and a single entry, say for example of $\Sigma_{\eta u}$, is denoted as $\sigma_{\eta_i u_j}$ or $\rho_{\eta_i u_j}\sigma_{\eta_i}\sigma_{u_j}$ for each $i, j \in \{1, 2, 3\}$. Finally, the set of parameters governing the joint dynamics of returns and predictors will be denoted as ϕ.

Our predictive system can be viewed as a reduced-form model that is consistent with a broad range of economic frameworks, rational or behavioral, where each return component r_i linearly depends on a lagged observable predictor x_i and a lagged unobservable predictor π_i. The predictability literature generally assumes that observable predictors evolve gradually over time and follow a first-order autoregressive process (e.g., Stambaugh, 1999). While this assumption is routinely used in many applications and observable predictors can be an important source of predictability, their slow evolution makes it challenging to precisely capture the true expected return (e.g., Pástor and Stambaugh, 2009). As a result, one may underestimate the uncertainty faced by an investor assessing the variance of future expected returns. To address the imperfect nature of observable predictors, we augment the predictive system with a driftless unobservable predictor, thus enhancing the ability of our predictive system to capture variations in expected returns compared to standard predictive regressions.

2.3 Multiperiod Predictive Variance

Our objective is to study the predictive variance over long horizons of rx_{t+1} and how the shape of the variance curve is affected by its underlying return components. Unlike the (ex-post) realized variance, which implicitly assumes full knowledge of the data generating process, the (ex-ante) predictive variance only conditions on information available to an investor and incorporates parameter uncertainty to make forward-looking predictions.

Define the k-period excess return in Equation (1) from period T through period $T+k$ as

$$rx_T^k = \sum_{i=1}^{3} r_{i,T}^k \tag{6}$$

with $r_{i,T}^k = \sum_{\ell=1}^{k} r_{i,T+\ell}$ for each $i = \{1, 2, 3\}$, i.e., the k-period foreign bond excess return, real interest rate differential, and real exchange rate return between T and $T+k$. Let D_T denote the information set available to an investor at time T, which includes past returns and observable predictors, while excluding any information on the unobservable predictors and the set of parameters governing the joint dynamics of returns and predictors. These elements are considered random, given that they are unknown to an investor. We can then assess the multiperiod predictive variance of excess returns given the information set available at time T as

$$\mathbb{V}\text{ar}\left(rx_T^k \mid D_T\right) = \sum_{i=1}^{3}\sum_{j=1}^{3} \mathbb{C}\text{ov}(r_{i,T}^k, r_{j,T}^k \mid D_T), \tag{7}$$

which comprises the multiperiod predictive variances and covariances of foreign bond excess returns, real interest rate differentials, and real exchange rate returns. In our setting, an investor is uncertain about the unobservable predictors π_T and the parameters ϕ of the predictive system. This allows us to decompose the multiperiod predictive variance as

$$
\begin{aligned}
\mathbb{V}\mathrm{ar}\left(rx_T^k \mid D_T\right) &= \sum_{i=1}^{3}\sum_{j=1}^{3}\mathbb{E}\left[\mathbb{C}\mathrm{ov}(r_{i,T}^k, r_{j,T}^k \mid \pi_T, \phi, D_T) \mid D_T\right] \\
&+ \sum_{i=1}^{3}\sum_{j=1}^{3}\mathbb{C}\mathrm{ov}\left[\mathbb{E}(r_{i,T}^k \mid \pi_T, \phi, D_T),\, \mathbb{E}(r_{j,T}^k \mid \pi_T, \phi, D_T) \mid D_T\right]. \qquad (8)
\end{aligned}
$$

The first term on the right side of this decomposition denotes the expectation of the conditional covariance of k-period returns, whereas the second term indicates the covariance of the conditional expectation of k-period returns. While an investor with a knowledge of π_T and ϕ only cares about the first term, an investor uncertain about π_T and ϕ also accounts for the second term. As a result, an international bond investor may perceive her excess returns disproportionately more volatile at long horizons. We now move to decompose both the expected conditional covariance and the covariance of expected returns on the right side of Equation (7).

2.4 Expected Conditional Covariance

The expectation of the conditional covariance of k-period returns is an important building block of Equation (8). This expectation is taken with respect to the investor's information set D_T, and reflects both parameter uncertainty and uncertainty about the current expected return. This happens as the investor ignores both ϕ and π_T.

Proposition 1. *Assuming Equations (2)–(5) hold, the expectation of the conditional covariance of k-period returns is given by*

$$
\mathbb{E}\left[\mathbb{C}\mathrm{ov}(r_{i,T}^k, r_{j,T}^k \mid \pi_T, \phi, D_T) \mid D_T\right] = \mathbb{S}_1(k) + \mathbb{S}_2(k) + \mathbb{S}_3(k), \qquad (9)
$$

where the terms on the right side represent three sources of uncertainty defined as

$$\mathbb{S}_1(k) = \underbrace{\mathbb{E}\left\{k\sigma_{u_i}\sigma_{u_j}\rho_{u_iu_j} \mid D_T\right\}}_{\textit{iid uncertainty}}, \tag{10}$$

$$\begin{aligned}\mathbb{S}_2(k) &= \underbrace{\mathbb{E}\left\{k\sigma_{u_i}\sigma_{u_j}[b_i\bar{e}_i\rho_{v_iu_j}A_{\gamma_i}(k) + b_j\bar{e}_j\rho_{u_iv_j}A_{\gamma_j}(k)] \mid D_T\right\}}_{\textit{mean reversion } x} \\ &+ \underbrace{\mathbb{E}\left\{k\sigma_{u_i}\sigma_{u_j}[\bar{d}_i\rho_{\eta_iu_j}A_{\delta_i}(k) + \bar{d}_j\rho_{u_i\eta_j}A_{\delta_j}(k)] \mid D_T\right\}}_{\textit{mean reversion } \pi}, \end{aligned}\tag{11}$$

$$\begin{aligned}\mathbb{S}_3(k) &= \underbrace{\mathbb{E}\left\{k\sigma_{u_i}\sigma_{u_j}b_ib_j\bar{e}_i\bar{e}_j\rho_{v_iv_j}B_{\gamma_i\gamma_j}(k) \mid D_T\right\}}_{\textit{future uncertainty } x} \\ &+ \underbrace{\mathbb{E}\left\{k\sigma_{u_i}\sigma_{u_j}\bar{d}_i\bar{d}_j\rho_{\eta_i\eta_j}B_{\delta_i\delta_j}(k) \mid D_T\right\}}_{\textit{future uncertainty } \pi} \\ &+ \underbrace{\mathbb{E}\left\{k\sigma_{u_i}\sigma_{u_j}[b_i\bar{e}_i\bar{d}_j\rho_{v_i\eta_j}B_{\gamma_i\delta_j}(k) + b_j\bar{d}_i\bar{e}_j\rho_{\eta_iv_j}B_{\delta_i\gamma_j}(k)] \mid D_T\right\}}_{\textit{future uncertainty } x \textit{ and } \pi}. \end{aligned}\tag{12}$$

The quantities $\bar{e}_s$ and $\bar{d}_s$ for $s = \{i, j\}$, $A_\chi(k)$ for $\chi = \{\gamma_i, \gamma_j, \delta_i, \delta_j\}$, and $B_{\chi\psi}(k)$ for $\chi\psi = \{\gamma_i\gamma_j, \delta_i\delta_j, \gamma_i\delta_j, \delta_i\gamma_j\}$ are functions of the parameters underlying the predictive system.

Proof. See Internet Appendix A.1 □

According to Proposition 1, the expectation of the conditional covariance of k-period returns consists of three distinct sources of uncertainty. The first component in Equation (10), denoted as $\mathbb{S}_1(k)$, can be interpreted as the uncertainty arising from *iid* shocks. Depending on the sign of $\rho_{u_iu_j}$, the correlation between unexpected returns, this source of uncertainty will make a constant contribution per period at all investment horizons.

The second component in Equation (11), labelled as $\mathbb{S}_2(k)$, captures the uncertainty associated with mean reversion in expected returns, and comprises two terms. The first term captures the correlation between unexpected returns and shocks to observable predictors through the cross-correlation terms $\rho_{v_iu_j}$ and $\rho_{u_iv_j}$, whereas the second term represents the correlation between unexpected returns and shocks to unobservable predictors by means of the cross-correlation components $\rho_{\eta_i u_j}$ and $\rho_{u_i \eta_j}$. When these correlations are negative, the second source of uncertainty will have a negative contribution to the predictive covariance, thus reflecting mean reversion in expected returns. The literature on stock returns generally finds a negative correlation between unexpected returns and shocks to expected returns and concludes that stocks have lower per-period variance and are less risky for long-horizon investors (e.g., Campbell, 1991; Campbell, Chan and Viceira, 2003).

The third component in Equation (12), named $\mathbb{S}_3(k)$, indicates the uncertainty about future expected returns, and comprises three terms. The first term captures the uncertainty about the future values of the observable predictors, while the second term reflects the uncertainty about future values of the unobserved predictors. The last term, instead, captures the joint uncertainty about the future values of observable and unobserved predictors. Even with full knowledge of the predictive system's parameters and the current value of the unobservable predictors, an investor is still uncertain about the future expected returns in each period. As a result, the third source of uncertainty produces additional predictive covariance, especially in the more distant future periods, that is often ignored in the asset pricing literature (e.g., Pástor and Stambaugh, 2012).

When returns are unpredictable, $\mathbb{S}_1(k)$ is the only source of uncertainty that matters in Equation (13). When returns are predictable, using observable and/or unobservable predictors, $\mathbb{S}_1(k)$ continues to remain the only source of uncertainty only for $k = 1$, since $A_x(1)$ and $B_{x,y}(1)$ are both equal to zero. As k grows, both $\mathbb{S}_2(k)$ and $\mathbb{S}_3(k)$ become increasingly important, since $A_x(k)$ and $B_{x,y}(k)$ are strictly increasing from zero to one as k moves from one to infinity. Hence, when returns are predictable and $k > 1$, all three sources of uncertainty play a role in Equation (13). While $\mathbb{S}_2(k)$ is likely to have a damping effect given past evidence on the correlation between unexpected returns and shocks to predictors, $\mathbb{S}_3(k)$ is likely to have an amplifying impact as the investor ignores the future values of expected returns. Finally, without observable predictors, all terms involving b_i and b_j in $\mathbb{S}_2(k)$ and $\mathbb{S}_3(k)$ will vanish. Similarly, without unobservable predictors, all terms involving δ_i and δ_j in $\mathbb{S}_2(k)$ and $\mathbb{S}_3(k)$

will disappear.[3]

To sum up, using the estimates of the predictive system in Equations (2)–(5), we can compute the expectation of the conditional covariances of the k-period returns through closed-form solutions. This expectation includes three different terms, namely, a first term reflecting *iid uncertainty* arising from a common assumption about returns, a second term capturing *mean reversion uncertainty* arising from both observable and unobservable predictor, and a third term referring to *future uncertainty about expected returns* stemming from the investor ignoring the future values of the observable predictor, the future values of the unobservable predictor, and their joint future dynamics. Taken together, an investor must consider all these different layers of uncertainty applied to three different return components, to better understand the multi-period predictive variance associated with her international bond strategy.

2.5 Covariance of Expected Returns

The second term on the right side of Equation (7) is the covariance of the conditional expected k-period returns given the investor's information set D_T.

Proposition 2. *Assuming Equations (2)–(5) hold, the covariance of the conditional expected k-period returns is given by*

$$\mathbb{C}\mathrm{ov}\left[\mathbb{E}(r_{i,T}^{k} \mid \pi_T, \phi, D_T),\, \mathbb{E}(r_{j,T}^{k} \mid \pi_T, \phi, D_T) \mid D_T\right] = \mathbb{S}_4(k) + \mathbb{S}_5(k), \tag{13}$$

where the terms on the right side represent two sources of uncertainty defined as

$$\mathbb{S}_4(k) \;=\; \underbrace{\mathbb{E}\left\{\frac{1-\delta_i^k}{1-\delta_i}\frac{1-\delta_j^k}{1-\delta_j} q_{ij,T} \mid D_T\right\}}_{\textit{current uncertainty } \pi}, \tag{14}$$

[3] When $i = j$, there are no observable predictors, and there is no uncertainty about ϕ, the expectation of the conditional covariance in Equation (13) coincides with the conditional variance presented in Pástor and Stambaugh (2012). See their Equation (6) on page 438.

$$\mathbb{S}_5(k) = \underbrace{\mathbb{C}\text{ov}\left\{k\mathbb{E}_{r_i} + \frac{1-\gamma_i^k}{1-\gamma_i}b_{i,T} + \frac{1-\delta_i^k}{1-\delta_i}c_{i,T},\, k\mathbb{E}_{r_j} + \frac{1-\gamma_j^k}{1-\gamma_j}b_{j,T} + \frac{1-\delta_j^k}{1-\delta_j}c_{j,T} \mid D_T\right\}}_{\textit{estimation risk}}. \tag{15}$$

The quantities $c_{s,T}$ and $b_{s,T}$ for $s = \{i, j\}$, $\mathbb{E}_x$ for $x = \{r_i, r_j\}$, and $q_{ij,T}$ are functions of the parameters underlying the predictive system.

Proof. See Internet Appendix A.2 □

According to Proposition 2, the covariance of the conditional expected k-period returns consists of two different sources of uncertainty. The first component in Equation (14), denoted as $\mathbb{S}_4(k)$, accounts for the investor's uncertainty about the current value of unobservable predictors affecting expected returns. It depends on the conditional covariance between the unobservable predictors $\pi_{i,T}$ and $\pi_{j,T}$ through $q_{ij,T}$ as well as their level of persistence captured by δ_i and δ_j, respectively. Overall, this term can play an important role as it reflects information not directly available to the investor at time T due to predictor imperfection.

The second component in Equation (15), named $\mathbb{S}_5(k)$, reflects estimation risk arising from uncertainty about ϕ, i.e., the parameters of the predictive system may not be perfectly estimated and thus affect the accuracy of predictions for future returns. This component combines the unconditional expected returns ($\mathbb{E}_{r_i}$ and $\mathbb{E}_{r_j}$) with the conditional means of observable ($b_{i,T}$ and $b_{j,T}$) and unobservable ($c_{i,T}$ and $c_{j,T}$) factors, respectively. While the previous four sources of uncertainty are expectations of random quantities due to uncertainty about ϕ, this component is the covariance of quantities whose randomness is also due to parameter uncertainty. In the absence of such uncertainty, this component is zero, which is why we attribute it the interpretation of estimation risk akin to Pástor and Stambaugh (2009).

In summary, the covariance of the conditional expected k-period returns captures both the immediate impact of unobservable predictors and the estimation risk associated with parameter uncertainties, each contributing to a more comprehensive risk assessment over varying time horizons.

2.6 Predictive Variance Ratio

Our goal is to calculate the multiperiod predictive variance outlined in Equation (7). This variance includes three predictive variances specific to foreign bond excess returns, real interest rate differentials, and real exchange rate returns, and three predictive covariances, each counted twice, involving the interactions between foreign bond excess returns and real interest rate differentials, foreign bond excess returns and real exchange rate returns, and real interest rate differentials and real exchange rate returns. Each of these components further incorporates five distinct sources of uncertainty detailed in Proposition 1 and Proposition 2. In our empirical analysis, we will evaluate the importance of these components and how they vary with the investment horizon k by presenting, similar to Pástor and Stambaugh (2009), the k-period variance ratio defined as

$$\mathbb{VR}(k) = \frac{\mathbb{V}\text{ar}\left(rx_T^k \mid D_T\right)}{k\,\mathbb{V}\text{ar}\left(rx_T^1 \mid D_T\right)} \tag{16}$$

$$= \sum_{i=1}^{3}\sum_{j=1}^{3} \frac{\mathbb{C}\text{ov}(r_{i,T}^k, r_{j,T}^k \mid D_T)}{k\,\mathbb{V}\text{ar}\left(rx_T^1 \mid D_T\right)}, \tag{17}$$

where $\mathbb{V}\text{ar}\left(rx_T^1 \mid D_T\right)$ is the one-month predictive variance of the international bond strategy. Throughout our empirical analysis, we will also scale the underlying components by the same amount to facilitate comparison across different horizons.

3 Data and Preliminary Statistics

This section describes the long-span dataset to estimate the long-horizon predictive variance of an investment strategy in which a US investor buys a constant maturity long-term bond in foreign currency while borrowing at the short-term deposit rate in domestic currency.

3.1 Long-Span Data

We focus on a sample of major countries relative to the US that exhibit a good degree of homogeneity and have relatively liquid and developed bond markets, such as Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, and the UK. The main source of our dataset is Global Financial Data, and the sample ranges between December 1799 and December 2023. The starting date varies across countries, depending on data availability, from December 1799 for the UK to December 1922 for New Zealand. We use the US as the domestic country and any other country as the foreign country.

The return vector r_{t+1} on the left side of Equation (2) includes the bond excess return in foreign currency, the short-term real interest rate differential between the foreign and domestic countries, and the real exchange rate return vis-á-vis the domestic currency. The bond excess return is calculated as $r_{1,t+1} = y^{\star}_{t+1} - i^{\star}_{t+1}$, where $y^{\star}_{t+1}$ is the log return on the *10-year total return government bond index* and $i^{\star}_{t+1}$ is the log return on the *total return bills index*, both defined between months t and $t+1$ for the same foreign country. The short-term real interest rate differential is computed as $r_{2,t+1} = (i^{\star}_{t+1} - \rho^{\star}_{t+1}) - (i_{t+1} - \rho_{t+1})$, where $i^{\star}_{t+1}$ and i_{t+1} are the log returns on the foreign and domestic *total return bills index* between months t and $t+1$, respectively, while $\rho^{\star}_{t+1}$ and ρ_{t+1} are the year-on-year log changes in the foreign and domestic *consumer price index* at time $t+1$, respectively. The *total return bills index* is largely based on 3-month treasury bill yields and deposit rates. When monthly data on the *consumer price index* are not available, we retrieve monthly observations by forward-filling lower-frequency data, i.e., we keep the last observation until a new one is made available. Finally, the real exchange rate return is constructed as $r_{3,t+1} = e_{t+1} + \rho^{\star}_{t+1} - \rho_{t+1}$, where e_{t+1} is the log return on the *spot exchange rate* between months t and $t+1$. The exchange rate is defined in units of US dollars per unit of foreign currency, so that an increase indicates foreign currency appreciation.

The set of observed predictors x_t on the right side of Equation (2) includes the term spread, the real output growth differential, and the short-term interest rate differential, due to data availability. The term spread is measured as the difference between the *10-year government bond yield* and the *3-month treasury bill yield*. When the latter is not directly available, we use the log return on the total return bills index as a proxy. The real output growth differential is quantified as the difference between the foreign and domestic year-on-year log changes in *real GDP*. We retrieve monthly data from quarterly or annual data via forward filling. Finally, the short-term interest rate differential is computed as the difference between the foreign and domestic *3-month treasury bill yield*.

3.2 Summary Statistics

We present the descriptive statistics for both returns and observed predictors in Table 1. The bond excess return in local currency is generally positive across countries and ranges between 0.26 and 1.54 percent per annum for the UK and Australia, respectively. The exception is New Zealand for which the average bond excess return is negative and equal to -0.51 percent per annum. The standard deviation, reported in percentage per annum, is fairly large and goes from 8.42 for the UK to 5.12 for Switzerland. The short-term real interest rate differential goes from -0.43 percent per annum for Germany to 1.69 percent per annum for Canada but the evidence suggests that the short-term real interest rate for the US has been, on average, lower than the short-term real interest rate abroad in our sample. The cross-sectional variation as measured by the standard deviation is fairly small and we record the largest value of 2.27 percent per annum for Japan. The real exchange rate evolves around zero, being negative for six countries and positive for the remaining three countries. We record the largest negative return for Australia (i.e., -0.72 percent per annum) and the largest positive value for Switzerland (i.e., 0.73 percent per annum). The standard deviation of the real exchange rate return, moreover, is generally larger than the standard deviation reported for real interest rate differentials. For example, Australia displays a standard deviation of 8.45 percent per annum, whereas New Zealand has a standard deviation of 11.69 percent per annum.

Table 1 about here

In addition to sample averages and standard deviations, we also measure the sample higher moments. The conventional measures of skewness and kurtosis, however, can be arbitrarily large especially when the sample is contaminated by large values. This is the case when one works with long-span samples. Our sample is indeed characterized by a number of events, including financial crisis, different monetary policy regimes as well as periods with fixed and floating exchange regimes. Instead of manually removing large values, we use robust measures of skewness and kurtosis (e.g., Kim and White, 2004). In particular, the coefficient of skewness is defined as $skew = (\mu - Q_2)/\sigma$, where μ is the sample mean, Q_2 is the sample mode, and σ is the sample standard deviation. The centred coefficient of kurtosis, moreover, is computed as $kurt = ((F^{-1}_{0.975} + F^{-1}_{0.025})/(F^{-1}_{0.75} + F^{-1}_{0.25})) - 2.91$, where F^{-1} denotes the quantile of the empirical distribution. We find that the coefficient of skewness is by and large negative but small in size. In contrast, the coefficient of kurtosis is sizeable, especially for the real exchange rate return. We also compute the first-order serial correlation. We find that bond excess returns and real exchange rate returns have a relatively low coefficient of serial correlation, whereas real interest rate differentials display a very high coefficient.

Table 1 also reports descriptive statistics of the observable predictors used to forecast bond excess returns, real interest rate differentials and real exchange returns. The mean of term spreads is overall positive, denoting upward sloping yield curves for the majority of the countries. The mean of real output growth differentials is generally negative (except for Canada and Japan), suggesting that output growth in the US has been, on average, higher than output growth in other countries. Finally, the mean of the nominal interest rate differential is positive for most countries (except for Germany and Switzerland). As already reported in existing studies, all three predictors exhibit a very high degree of persistence.[4]

[4]We remove the hyperinflation period between January 1919 and December 1924 for Germany and the post-WWII period between January 1945 and July 1948 for Japan so that our analysis is not affected by extremely large data points.

4 Empirical Results

In this section, we present our findings on the estimation of the long-horizon predictive variance. First, we present the parameter estimates for the predictive system described earlier, and then illustrate the predictive variance and its underlying components for the baseline case involving the UK relative to the US, given the extended availability of historical data. We then extend our focus to all other countries in our sample and evaluate the consistency of our key findings. Finally, we check whether the observed patterns are unique to recent years or have persisted over the past century.

4.1 Benchmark Priors

We use Bayesian methods to estimate the parameters of the predictive system in Equations (2)–(5). Similar to Pástor and Stambaugh (2009), the priors are assumed to be independent across parameters and follow conventional functional forms such as Normal and Wishart distributions. For each parameter, we specify a benchmark prior for our core analysis as well as two alternative priors in the robustness analysis reported in Section 5.1. When we deviate from the benchmark prior for a given parameter, we hold the benchmark priors for all other parameters. We estimate the predictive system under each specification and explore the extent to which the long-horizon variance is sensitive to prior beliefs. We start by reviewing the benchmark priors before moving the parameter estimation.

The benchmark priors on a_i and b_i, the constant and slope coefficient on the observable predictor in Equation (2), are normally distributed. The prior means are obtained by estimating the regressions $r_{i,t+1} = a_i + b_i x_{i,t} + u_{i,t+1}$ via least-squares, while the prior variances are fixed at 0.5. The benchmark priors on θ_i and γ_i, the constant and autoregressive coefficient on the observable predictor in Equation (3), are also normally distributed. The prior means are set by estimating the regressions $x_{i,t+1} = a_i + b_i x_{i,t} + v_{i,t+1}$ via least squares, with prior variances set to 0.5 for θ_i and 0.01 for γ_i. The benchmark prior on δ_i, the autoregressive coefficient on the unobservable predictor in Equation (4) is also normally distributed, with prior mean of 0.99 and a prior variance of 0.01. Additionally, the prior distribution on γ_i is restricted to the interval $(-1, 1)$, whereas the prior distribution in δ_i is restricted to the interval $(0, 1)$.

The covariance matrices Σ_{uu}, Σ_{vv}, and Σ_{vu} in Equation (5) are stack together to form $\Sigma_{\varepsilon\varepsilon}$ and then sampled jointly. The benchmark prior on $\Sigma_{\varepsilon\varepsilon}^{-1}$ is specified as a diagonal Wishart distribution, with prior parameters chosen so that the mean of diagonal elements matches the sample variances of r_t and x_t. The benchmark prior on $\Sigma_{\eta\eta}^{-1}$ is also specified as a diagonal Wishart distribution, with prior parameters chosen so that the prior mean of the diagonal elements is one. The benchmark prior on $\Sigma_{u\eta}$ follows a Normal distribution, and we set the prior parameters so that the implied prior on $\rho_{u_i\eta_i}$, the correlation between unexpected returns and shocks to the unobservable predictor, has a mean of -0.5 and 95% of its probability mass lies within the interval $[-0.75, 0.25]$. This prior follows the argument of Pástor and Stambaugh (2009), who suggest that, a priori, the correlation between unexpected return and the innovation in expected return is likely to be negative. Finally, the benchmark prior on Σ_{vu} follows a Normal distribution with zero means and unit variances. Additional details on the priors are reported in Internet Appendix B.4.

4.2 Parameter Estimation

Armed with long-span data and benchmark priors, we estimate the parameters of the predictive system using a Gibbs sampling algorithm and incorporating the *forward filtering, backward sampling* method of Carter and Kohn (1994) to sample the vector of unobservable predictors. The algorithm runs for 100,000 iterations, following an initial burn-in period of 20,000 iterations, and we retain one in every ten iterations to mitigate potential serial correlation. We then compute posterior means, standard deviations (STD), numerical standard errors (NSE), relative numerical inefficiency (RNI), and highest posterior density intervals to assess statistical significance. A detailed description of our algorithm can be found in Internet Appendix B.

In Table 2, we report the posterior estimates associated with Equation (2). Recall that r_{t+1} includes the bond excess return in local currency, the real interest rate differential between the foreign country and the US, the real exchange rate return relative to the US dollar, while x_t comprises the term spread, the real output differential between the foreign country and the US, and the nominal interest rate differential between the foreign country and the US. We use the superscripts *, **, and *** to indicate that the 90%, 95%, and 99% highest posterior density intervals, respectively, do not contain zero.

Table 2 about here

The estimates of b_1 suggest that bond excess returns are predictable by the lagged local-currency term spreads in all cases. For most countries, the significance is at the 1 percent level using highest posterior density intervals, corroborating and extending the evidence already reported in the existing literature for the US Treasury bond market (e.g., Joslin et al., 2014). The recorded R^2 are also consistent with the ones recorded for the US at a monthly frequency (Gargano, Riddiough and Sarno, 2018). Similar evidence of predictability is found for b_2, the slope associated to the predictability of the real interest rate differential using the lagged real output differential. All estimates of b_2 are statistically significant at the 1 percent level, with large R^2, all exceeding 92 percent. The evidence on the predictability of the real exchange rate returns is less pervasive as, consistently with much empirical evidence, the estimates of b_3 are only significant in 3 cases out of 9. Although there is compelling evidence of predictability for bond excess returns in local currency and real interest rate differentials, the predictive variance of the returns associated to our international bond strategy is likely to be affected by considerable uncertainty as real exchange returns turn out to be difficult to predict. In addition to parameter estimates, we also report the NSE and RNI (e.g., Della Corte, Sarno and Tsiakas, 2008b). The former is a measure of numerical precision and reflects the variability in the posterior mean estimate if the simulation were repeated multiple times. Our NSE estimates are fairly small, suggesting that the number of draws to calculate our posterior moments is more than sufficient. The latter quantifies the efficiency loss when calculating the posterior mean from autocorrelated samples, compared to independent samples. Our RNI estimates are generally small, indicating that the draws for the computation of the posterior moments show little serial correlation. This is also due to the fact that we retain one in every ten iterations before calculating our posterior moments.

Table 3 about here

In Table 3, we report the posterior estimates associated to Equation (3). The evidence confirms the high persistence in the observable predictors for all countries. In fact, the estimates of γ_i are all statistically significant at the 1 percent level and their value exceeds 0.93 across countries and predictors. Moreover, the small numbers for NSE and RNI estimates confirms that we have used an adequate number of draws to construct out posterior moments.

Table 4 about here

Interestingly, the high persistence recorded for the observable predictors is mimicked by a similar persistence for unobservable predictors. Although the coefficients for δ_i reported in Table 4 are slightly smaller than the ones reported in Table 3, their magnitude is still substantially large. This suggests that the predictable, potentially slow-moving, part of bond excess returns, real interest rate differentials and real exchange rates require multiple persistent predictors to be captured. Given the long-sample period investigated, it is reasonable to hypothesize that unobserved predictors may be associated with abrupt yet persistent changes in monetary and exchange rate regimes that took place during the past two centuries. Internet Appendix Table A.1 plots the posterior mean of the unobservable predictors, and unveil some notable shifts. For example, the unobservable predictors exhibit large spikes around the Great Depression, Oil Crisis, Volcker Shock, Black Wednesday, and Global Financial Crisis, and Covid-19 Pandemic. We now move to assess the multiperiod predictive variance of our international bond strategy.

4.3 Predictive Variance: Baseline Case

Using the posterior draws of the parameters from the predictive system in Equations (2)–(5), we project the predictive variance over the next k periods, applying the closed-form solutions from Proposition 1 and Proposition 2. Specifically, D_T, the information set available to an investor at time T, encompasses all data up to December 2023, allowing us to compute the predictive variance from T to $T+k$, where k represents the future investment horizon. For the sake of exposition, our core analysis focuses on the UK relative to the US due to the extensive data availability. We subsequently assess other major countries to confirm the robustness of our main findings. To facilitate comparison across countries, we focus on the predictive variance ratio.

FIGURE 1 ABOUT HERE

In Figure 1, we start by plotting the predictive variance ratio of the international bond strategy as well as its constituents based on Equation (17) Panel A focuses on the total predictive variance ratio and reveals an upward pattern. The longer the investment horizon, the larger the predictive variance. Over a 20-year horizon, the monthly predictive variance of our strategy increases by 100 percent, exceeding 200 over a 50-year horizon. This finding is comparable to the long-horizon predictive variance of a US equity buy and hold strategy computed over a similar sample period by Pástor and Stambaugh (2012). This result also echoes early findings by Campbell and Viceira (2001, 2002) showing that long-term bonds are not very different from equities over longer investment horizons.

To gain insight on the key drivers behind our findings, we further examine the individual components of our strategy, i.e., the foreign bond excess return, the real interest rate differential, and the real exchange rate return. In Panel B, we illustrate the contribution of each component to the overall predictive variance by summing their respective variances and covariances. For instance, the contribution coming from the real exchange rate return includes its variance and its covariances with the bond excess return and the real interest rate differential. Panel B reveals that the upward trend observed in Panel A is largely driven by the real exchange rate return and the real interest rate differential, while the effect of the bond excess return in local currency diminishes with longer investment horizons.

Finally, in Panels C and D, we display the individual predictive variances and the covariances appearing on the right side of Equation (17). In light of our assessment, the upward pattern exhibited by the predictive variance of the strategy is mostly due to the individual predictive variances of the real exchange rate return and the real interest rate differential, with the former playing a more prominent role. The predictive variance of bond excess returns in local currency does not play a substantial role, as its impact is relatively constant across investment horizons. Among the three predictive covariances, those involving the real exchange rate returns are large and negative in sign. The results also suggest that the predictive co-movements between interest rates and real exchange rate are important in the long-run, as they tend to reduce the overall expected risk of the strategy, especially at longer horizons. This negative sign is also in line with the recent evidence reported in Engel (2016), whereby the correlation between long-term expected risk premia and real interest rate differentials is negative. This finding suggests that predictive co-movement between foreign bond excess returns and real interest rate differentials are less important in determining the long-term risk profile of the strategy.

Figure 2 about here

As shown in Equation (7), we can decompose the predictive variance into two parts: the expectation of the conditional covariance of k-period returns and the covariance of the conditional expectation of k-period returns. Panel A of Figure 2 illustrates this decomposition, indicating that the latter component is the main driver of the upward trend of the total predictive variance, while the former has a minimal impact, likely due to the extensive 200-year span of monthly data used in our analysis. Moreover, Proposition 2 shows that the expected conditional covariance consists of three source of uncertainty, i.e., *iid uncertainty*, *mean reversion* and *future expected return uncertainty*, while Proposition 1 unpacks the covariance of expected returns into two further sources of uncertainty, i.e., *current expected* *return uncertainty* and *estimation risk* via Proposition 2. Panel B of Figure 2 reports this decomposition, always in terms of variance ratio. The patterns exhibited by this decomposition are rather striking and they unambiguously assign a dominant, if not an exclusive role, to the uncertainty about future returns. This type of uncertainty is found to be important for long horizons in the context of US equity markets (e.g., Pástor and Stambaugh, 2012). However, its impact on the predictive variance of returns is partly offset by the effect of mean reversion in returns, especially at short to medium horizons. Panel C explores mean reversion, showing both observable and unobservable predictors follow a downward trend. Finally, in Panel D, we break down future expected return uncertainty by examining both observable and unobservable predictors and their interaction. The unobservable predictors dominate, as their impact on predictive variance over the long run is three times greater than that of observable predictors, while their interaction is negligible at very long horizons. If we think of unobservable predictors as variables associated with changes in monetary and exchange rate regimes that have occurred over the past two centuries and are not already captured by the set of observable predictors, the evidence suggests that these shifts, particularly their unpredictable effects on future returns, have a significant impact on the overall risk profile of the strategy.

4.4 Predictive Variance: Cross-Country View

To validate our main findings regarding the impact of future uncertainty, we extend our analysis to include all major countries, including the UK. Figure 3 presents the average value and interquartile range of the total predictive variance ratio, along with its underlying uncertainty components derived from Proposition 2 and Proposition 1. We observe qualitatively similar trends, with a consistent upward trajectory in the predictive variance and a significant contribution from future expected return uncertainty as the primary driver.

FIGURE 3 ABOUT HERE

In Figure 4, we further examine the components of international bond strategies, highlighting that real exchange rate returns and real interest rate differentials are the primary contributors to the risk profile of an international investor. In contrast, foreign bond excess returns play a comparatively minor role in shaping this profile.

Figure 4 about here

This cross-country decomposition emphasizes the robustness of our findings, highlighting the persistent role of future expected return uncertainty in shaping the predictive variance profile for international bond strategies. The results suggest that, regardless of the specific country, uncertainty tied to future expected returns remains a dominant factor driving the risk outlook over multiple periods, reinforcing the importance of forward-looking uncertainty in these strategies. Moreover, the consistent prominence of real exchange rate returns and real interest rate differentials across different countries indicates that exchange rate fluctuations and cross-country interest rate spreads are pivotal in determining the predictive variance faced by international investors. By contrast, foreign bond excess returns contribute less significantly. This comprehensive cross-country perspective therefore confirms that these factors are not isolated to a single country or currency regime but are fundamental to the international bond investment landscape as a whole.

4.5 Multiperiod Predictive Variance: An Expanding Window View

Our analysis leverages over 200 years of data to calculate the predictive variance of an international bond strategy projected over the next 50 years. A natural question arises: is the observed upward-sloping pattern a recent phenomenon, or has it persisted historically? To explore this, we re-estimate the parameters of the predictive model using an expanding window starting in December 1900 and updating annually until December 2023. For each set of parameters, we then compute the multiperiod predictive variance of the strategy for the next 50 years, and display the results in Figure 5 using a three-dimensional chart.

Figure 5 about here

The vertical axis represents the predictive variance, while the horizontal axes correspond to time periods and forecasting horizons. The color scale, moving from red (low) to yellow and then to blue (high), indicates the intensity of the predictive variance. Regions with more red signify lower variance, while blue regions signify higher variance. This chart shows a general upward slope in the predictive variance over time, in line with our previous findings. Interestingly, the predictive variance curve was steeper at the beginning of the last century, peaking around the end of World War I. It then gradually became less steep, maintaining a relatively stable slope up until the early 1980s. From that point onward, the curve begins to rise again, regaining its steepness and sustaining this upward trend through to the present. This pattern suggests that while economic and policy shocks initially drove high predictive variance, the relative stability mid-century was later disrupted, leading to renewed variability in recent decades.

Figure 6 about here

In Figure 6, we examine the predictive variances of the components underlying the international bond strategy. The predictive variance of the foreign bond excess return has generally remained flat, with notable exceptions during World War II, where the variance ratio surged to 2 for the next projected 50 years, and in the 1970s, where it reached 1.5. These spikes align with significant economic challenges. During WWII, UK government debt rose to over 200% of GDP as the war strained public finances. In the 1970s, the UK faced economic difficulties, often referred to as 'the sick man of Europe" due to sluggish growth and deteriorating industrial relations since the end of the war. These historical contexts help explain the periods of heightened predictive variance in the international bond strategy.

The predictive variance of the real interest rate differential was notably steep at the beginning of the last century, with the variance ratio reaching 5 for the next projected 50 years. Over time, this variance gradually declined, leveling out to around 2 for the same horizon in recent years. One possible explanation for this downward trend is the improvement in monetary policy and increased policy coordination, particularly following the collapse of the

Bretton Woods system. Enhanced central bank practices and coordinated economic policies likely contributed to stabilizing interest rate differentials, reducing the associated predictive variance in the long term.

The predictive variance of the exchange rate return has consistently shown an upward slope, with a variance ratio around 2 for projections 50 years ahead. This ratio dipped below 2 in the early 1980s, a period marked by the Federal Reserve's aggressive rate hikes under Paul Volcker to combat inflation. Conversely, it rose above 2 in the early 1990s, coinciding with the pound's depegging from the Deutsche Mark and the infamous speculative attack on the Bank of England by George Soros. These events highlight how major monetary policy shifts and currency market pressures have historically influenced the long-term predictive variance of the real exchange rate return.

Figure 7 about here

In Figure 7, we analyze the predictive covariances, each counted twice, among the components of the international bond strategy. These predictive covariances consistently exhibit a negative slope, reflecting opposing movements in the components over time. The predictive covariance between foreign bond excess returns and the real interest rate differential shows a pronounced negative slope up to the early 1980s, becoming less steep over the last 30 years. This shift may reflect increased stability in interest rate differentials and reduced volatility in bond returns. The predictive covariance between foreign bond excess returns and the real exchange rate return has also been negative, displaying regular fluctuations. This pattern of ups and downs suggests sensitivity to economic cycles and exchange rate adjustments, impacting the long-term predictive relationship between bond returns and exchange rate movements. The predictive covariance between the real interest rate differential and the real exchange rate return shows a very steep negative slope until the early 1970s, after which it becomes significantly less steep. This shift aligns with the collapse of the Bretton Woods system and the transition to floating exchange rates, which likely reduced the influence of tightly controlled exchange rate policies on interest rate differentials. The transition to a floating regime appears to have moderated the relationship between these two components, resulting in a less pronounced negative slope in recent decades.

FIGURE 8 ABOUT HERE

In Figure 8, we examine the uncertainty components underlying the international bond strategy. Our analysis confirms that future expected return uncertainty has consistently been the dominant driver of predictive variance. This component was especially high at the start of the last century, peaking around World War I, before gradually declining. Interestingly, it began to rise again in the early 1990s, reflecting renewed economic and financial uncertainty during that period. This pattern underscores the impact of historical events and economic shifts on the long-term predictive variance of an international bond investor.

5 Robustness

In this section, we conduct several robustness exercises to verify the stability of our findings. We begin by exploring the role of priors, evaluating how different prior specifications impact our predictive variance estimates. Second, we examine the role of unobservable predictors by estimating a predictive system without observable predictors. Finally, we assess the importance of predictor imperfection by only using observable predictors in our analysis.

5.1 Role of Priors

In our core analysis, we use benchmark priors as described in Section 4.2 to estimate the predictive variance of an international bond strategy involving the UK relative to the US. These priors are chosen to ensure that parameter estimates are grounded in observed data. To test the robustness of our findings, we explore for selected parameters two alternative priors, referred to as the flexible prior and the loose prior, that introduce greater variability.

By comparing the results under these alternative priors with those of the benchmark, we assess the sensitivity of the predictive variance to different prior specifications.

Figure 9 and Table 5 about here

We start with the slope coefficients in b. The benchmark prior has a mean based on actual data and a variance of 0.5. The flexible prior keeps the same mean but increases the variance to 1, while the loose prior has a mean of 0 and a variance of 2. Despite increased variability in the priors, Figure 9 shows that the posterior densities remain similar, indicating minimal impact on the estimates. Table 5 confirms that the predictive variance ratios are stable across priors, with values of 2.00, 2.05, and 2.07 at 20 years for the benchmark, flexible, and loose priors, respectively.

Figure 10 and Table 6 about here

For the autoregressive coefficients in γ, the benchmark prior has a mean derived from data and a variance of 0.01. The flexible and loose priors have means of 0.95 and 0.90, respectively, with the same variance of 0.01. As shown in Figure 10, the posterior densities remain largely unchanged, and Table 6 demonstrates stable predictive variance ratios, with virtually no differences at all horizons.

Figure 11 and Table 7 about here

Examining the autoregressive coefficients in δ, we observe similar results. The benchmark prior has a mean of 0.99 and a variance of 0.01, while the flexible and loose priors use means of 0.95 and 0.90 with the same variance. Figure 11 and Table 7 show that the posterior estimates and predictive variance ratios remain consistent across priors, with values of 2.00, 1.98, and 1.95 at 20 years for the benchmark, flexible, and loose priors, respectively.

Figure 12 and Table 8 about here

Finally, we analyze the correlations between unexpected returns and shocks to the unobservable predictor, $\rho_{u\eta}$. The benchmark prior has a mean of -0.5 with 95% of its probability mass in the range of $[-0.75, -0.25]$. The flexible prior widens this range to $[-1, 0]$, while the loose prior shifts to a mean of 0 with a range of $[-1, 1]$. Figure 12 indicates slight shifts in posterior densities for the foreign bond excess return, with stable results for other components. Table 8 confirms that the predictive variance ratios remain similar, highlighting the robustness of our findings across different prior settings. For example, with values of 2.00, 1.90, and 1.94 at 20 years for the benchmark, flexible, and loose priors, respectively.

5.2 Unobservable Predictors

In our core analysis, we use a predictive system that incorporates both observed and observable predictors, finding that the steepness of the predictive variance is driven by future expected return uncertainty arising from the real interest rate differential and real exchange rate return. To test the robustness of these findings, we re-estimate the predictive system using only the unobserved predictors in π_t, thus disregarding the influence of the observable predictors in x_t, and then recompute the predictive variance at different horizons.

Figure 13 about here

Specifically, we use the results from Proposition 1 and Proposition 2, while setting all parameters related to the observable predictors equal to zero. This adjustment eliminates *mean reversion in x*, *future uncertainty in x* and *future uncertainty in π and x*. The predictive variance ratio and its components, shown in Figure 13, reveal very similar results, supporting our main findings.

5.3 Observable Predictors

Predictor imperfection is an important element of our framework. As highlighted by Pástor and Stambaugh (2012), ignoring imperfections in predictors can lead to an underestimation of the uncertainty faced by investors. To assess whether this is the case in our setup, we re-estimate the predictive system using only the observable predictors in x_t, thus ignoring the influence of the unobservable predictors in π_t, and recompute the predictive variance at different horizons. By doing so, we can isolate the role of unobserved factors in contributing to predictive variance.

Figure 14 about here

For our calculations, we leverage on the results from Proposition 1 and Proposition 2, while setting all parameters related to the unobservable predictors equal to zero. This means that *mean reversion in* π, *future uncertainty in* π, *future uncertainty in* π *and* x, and *current uncertainty in* π all vanish. The results, shown in Figure 13, indicate a significant reduction in the predictive variance ratio when unobservable predictors are excluded. Specifically, the variance ratio increases roughly to 1.2 at a 50-year horizon, compared to about 2.3 when both observable and unobservable predictors are considered. This discrepancy highlights the importance of accounting for predictor imperfections, as they meaningfully contribute to the perceived risk over longer investment horizons.

5.4 Conditional Variance

Our analysis is centered around the predictive variance, which captures multiple sources of uncertainty as well as predictor imperfection. Naturally, one might question whether predictive variance differs from conditional variance. To address this, we compute the multiperiod conditional variance of our investment strategy, which excludes both parameter uncertainty and uncertainty about future returns. To perform this calculation, we begin by using the posterior means of parameter estimates derived from the predictive system that exclusively utilizes observable predictors. Hence, we compute the conditional variance by summing the *iid variance* and *mean reversion in* x. The latter corresponds to the term within the expectation in Equation (10), while the former term appears in the first expectation in Equation (11).

Figure 15 about here

We present the conditional variance in Figure 15, confirming the downward slope observed by Viceira and Wang (2018). For example, the conditional variance ratio, calculated by scaling the conditional variance of k-period excess returns by that of the one-period excess return, declines to approximately 0.82 at 10 years and 0.72 at 50 years. We also observe that the variance of the real interest rate differential and real exchange rate remains stable, as the covariances between real interest rate differentials with foreign bond excess returns and real exchange rate returns, respectively. Our findings suggest that by ignoring imperfections in predictors and uncertainties about the model's parameters and future returns, we are likely to underestimate the magnitude of underlying variances and covariances in the investment strategy. In other words, by leaving out these uncertainties in the conditional variance, we miss fluctuations that contribute to the overall predictive variance, resulting in a lower estimate than if we accounted for these elements. Put differently, while conditional variance gives a simpler view of risk, it fails to fully capture the range of possible variations in returns when we factor in predictor imperfection, model stability, and the uncertainty about future returns. Predictive variance, which includes both observable and unobservable predictors, therefore provides a more realistic measure of risk, especially over longer periods when these uncertainties play a larger role.

5.5 Floating Exchange Rate Regime

Another potential concern regarding our findings is that our long-span sample includes both the Gold Standard and Bretton Woods periods, during which exchange rates were pegged, potentially impacting the variability and behavior of real exchange rate returns. To address this, we conduct our analysis using data from the floating exchange rate period between January 1973 and December 2023.

Figure 16 about here

Figure 16 presents the predictive variance along with the underlying return components. The overall shape of the predictive variance curve remains consistent with our core findings, albeit slightly steeper. In this floating rate period, the real exchange rate return stands out as the primary contributor to predictive variance, while the real interest rate differential plays a negligible role. This may happen as the floating period is marked by the Great Moderation, which likely muted the role of the interest rate differential. Also, both the Federal Reserve and the Bank of England adopted a rules-based approach to monetary policy, which helped to anchor inflation expectations and stabilize the economy. Additionally, we observe a slightly positive covariance between the real interest rate differential and the real exchange rate return. This suggests a tendency for foreign currencies to appreciate when local interest rates rise relative to US domestic rates, reflecting a possible short-term response of capital flows to interest rate changes. We also show the decomposition of uncertainty. This analysis reaffirms our previous conclusions that future uncertainty remains a key driver of the predictive variance for the bond investment strategy. The dominance of this component underscores the role of unanticipated factors and structural changes, which persist as significant sources of risk in the floating exchange rate environment.

5.6 Investing in US Bonds

Our analysis takes the perspective of a US investor going long on a constant-maturity bond denominated in foreign currency while borrowing at the short-term rate in dollars. A natural extension is to reverse the exercise and consider, for example, a UK investor going long on a constant-maturity bond denominated in US dollars while borrowing at the short-term rate in pounds.

Figure 17 about here

We provide the outcome of this exercise in Figure 17, using the entire sample period spanning from 1799 to 2023. While the predictive variance continues to exhibit an upward-sloping pattern, it is less steep compared to our core analysis presented in Figure 1. As shown in

Panel A, the predictive variance ratio for a UK investor holding a US bond is approximately 1.5 at a 20-year horizon and rises to about 1.6 at a 50-year horizon. This occurs because the predictive variance of the real exchange rate is no longer the dominant component, as demonstrated in Panel C. This finding aligns with the fact that holding US dollars serves as a natural hedge during periods of market turmoil, given its tendency to strengthen as a safe-haven currency. Among the predictive covariances shown in Panel D, only the co-movements between real exchange rate returns and real interest rate differentials are significant. Similar to our core results, this covariance is predominantly negative at long horizons, consistent with the evidence in Engel (2016), which indicates a negative correlation between long-term expected risk premia and real interest rate differentials. Finally, as seen in Panel B, the dominant source of uncertainty for a UK investor is the future expected return uncertainty, which highlights the continued relevance of future return dynamics in shaping the predictive variance of an international bond strategy.

6 Conclusions

We consider a simple strategy where a US investor holds an unhedged position in a long-term foreign bond, funded at the domestic risk-free rate. The predictive variance on this strategy, which depends on foreign bond excess returns, real interest rate differentials between foreign and domestic country, and bilateral real exchange rate returns, can be decomposed into five distinct sources of uncertainty using closed-form solutions. Empirically, we use over 200 years of data for major countries relative to the US to assess the long-horizon risk associated with this strategy. To capture the complexities of the return dynamics, we account for imperfect predictability by allowing unobserved predictors to complement observable predictors in forecasting the different return components of the international bond strategy. This approach permits a comprehensive assessment of the uncertainty faced by an investor as noted by Pástor and Stambaugh (2009).

Our analysis reveals several important findings. First, the predictive variance of the investment strategy consistently increases with the investment horizon. This upward trend is primarily driven by the variances of both short-term interest rate differentials and real exchange rate returns. In contrast, the variance of the foreign bond excess return is less sensitive to the length of the investment horizon. We also examine the covariance components contributing to the risk profile of the bond strategy. We find that the predictive covariance between foreign bond excess returns and real exchange rate returns has a minimal influence on the long-term risk profile of the strategy. On the other hand, the predictive covariance between real interest rate differentials and real exchange rate returns plays an important role in reducing the overall expected risk, particularly at longer horizons. This negative covariance acts as a stabilizing force, helping to offset some of the individual uncertainties associated with these components.

Second, future uncertainty about expected returns plays a dominant role in shaping the predictive variance of the strategy at all investment horizons. While mean reversion could theoretically help to mitigate long-term risk, it fails to fully offset the future uncertainty about expected returns. This means that, even when bond returns, interest rates, and exchange rates exhibit signs of predictability, the inherent uncertainties are too large to be effectively countered by mean reversion alone.

Third, uncertainty about future returns is predominantly driven by components associated with unobserved predictors. This suggests that much of the uncertainty comes from factors that are not directly measurable or observable, reflecting the limitations of traditional predictors in fully capturing market dynamics. Moreover, the shape of the predictive variance over longer horizons can be interpreted as arising from shifts in monetary and exchange rate regimes that are not captured by the observable predictors. This suggests that unanticipated changes in policy and macroeconomic conditions play a critical role in shaping the long-term risk profile of the strategy.

References

Allen, Kate, "Austria Sells Record Largest €3.5bn Century Bond," *Financial Times*, 2017, *https://www.ft.com/content/1d4f47e2-97a7-11e7-b83c-9588e51488a0.*

Andrews, Spencer, Riccardo Colacito, Mariano Massimiliano Croce, and Federico Gavazzoni, "Concealed Carry," *Journal of Financial Economics*, 2024, *159*, 103874.

Avramov, Doron, Scott Cederburg, and Katarína Lučivjanská, "Are Stocks Riskier over the Long Run? Taking Cues from Economic Theory," *Review of Financial Studies*, 2018, *31*, 556–594.

Barberis, Nicholas, "Investing for the Long Run when Returns Are Predictable," *Journal of Finance*, 2000, *55*, 225–264.

Bogdanova, Bilyana, Tracy Chan, Kristina Micic, and Goetz von Peter, "Enhancing the BIS Government Bond Statistics," *BIS Quarterly Review*, 2021, *June*, 15–24.

Campbell, John Y., "A Variance Decomposition for Stock Returns," *Economic Journal*, 1991, *101*, 157–179.

_ and Luis M. Viceira, *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford: Oxford University Press, 2002.

_ and _ , "The Term Structure of the Risk-Return Tradeoff," *Financial Analysts Journal*, 2005, *61*, 34–44.

_ and Luis. Viceira, "Who Should Buy Long-Term Bonds?," *American Economic Review*, 2001, *91*, 99–127.

_ , Yeung Lewis Chan, and Luis M. Viceira, "A Multivariate Model of Strategic Asset Allocation," *Journal of Financial Economics*, 2003, *67*, 41–80.

Carter, C. K. and R. Kohn, "On Gibbs Sampling for State Space models," *Biometrika*, 1994, *81*, 541–553.

CGFS, "US Dollar Funding: An International Perspective," CGFS Paper No 65, Committee on the Global Financial System 2020.

Cochrane, John H. and Monika Piazzesi, "Bond Risk Premia," *American Economic Review*, 2005, *95*, 138–160.

Della Corte, Pasquale, Lucio Sarno, and Daniel L. Thornton, "The Expectation Hypothesis of the Term Structure of Very Short-Term Rates: Statistical Tests and Economic Value," *Journal of Financial Economics*, 2008, *89*, 158–174.

_ , _ , and Ilias Tsiakas, "An Economic Evaluation of Empirical Exchange Rate Models," *Review of Financial Studies*, 06 2008, *22* (9), 3491–3530.

Engel, Charles, "Exchange Rates, Interest Rates, and the Risk Premium," *American Economic Review*, 2016, *106*, 436–74.

_ **and Kenneth West**, "Exchange Rates and Fundamentals," *Journal of Political Economy*, 2005, *113*, 485–517.

Fama, Eugene and Robert R Bliss, "The Information in Long-Maturity Forward Rates," *American Economic Review*, 1987, *77*, 680–92.

Froot, Kenneth A., "Currency Hedging Over Long Horizons," *Annals of Economics and Finance*, 2019, *20* (1), 37–66.

Gargano, Antonio, Steven J. Riddiough, and Lucio Sarno, "Volume and Excess Returns in Foreign Exchange," *Working*, 2018, *Working Paper.*

Ilmanen, Antti, "Time-Varying Expected Returns in International Bond Markets," *Journal of Finance*, 1995, *50*, 481–506.

Joslin, Scott, Marcel Priebsch, and Kenneth J. Singleton, "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks," *Journal of Finance*, 2014, *69*, 1197–1233.

Kim, Tae-Hwan and Halbert White, "On More Robust Estimation of Skewness and Kurtosis," *Finance Research Letters*, 2004, *1*, 56–73.

Ludvigson, Sydney C. and Serena Ng, "Macro Factors in Bond Risk Premia," *Review of Financial Studies*, 2009, *22*, 5027–5067.

Meese, Richard and Kenneth Rogoff, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?," *Journal of International Economics*, 1983, *14*, 3–24.

Merton, Robert, "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case," *Review of Economics and Statistics*, 1969, *51*, 247–57.

Meyer, Josefin, Carmen M Reinhart, and Christoph Trebesch, "Sovereign Bonds Since Waterloo," *Quarterly Journal of Economics*, 2022, *137* (3), 1615–1680.

Pástor, Lubos and Robert F. Stambaugh, "Predictive Systems: Living with Imperfect Predictors," *Journal of Finance*, 2009, *64*, 1583–1628.

_ **and** _ , "Are Stocks Really Less Volatile in the Long Run?," *Journal of Finance*, 2012, *67*, 431–478.

Samuelson, Paul, "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 1969, *51*, 239–46.

Siegel, Jeremy J., "The Real Rate of Interest from 1800–1990: A Study of the U.S. and the U.K.," *Journal of Monetary Economics*, 1992, *29*, 227–252.

—, *Stocks for the Long Run*, New York: McGraw Hill, 2008.

Sindreu, Jon, "Investors Plow Into Long-Dated Bonds," *Wall Street Journal*, 2017, *March 31, Print Edition.*

Stambaugh, Robert F., "Predictive regressions," *Journal of Financial Economics*, 1999, *54*, 375–421.

Viceira, Luis. M. and Kevin Wang, "Global Portfolio Diversification for Long-Horizon Investors," *Working Paper*, 2018, *Harvard University.*

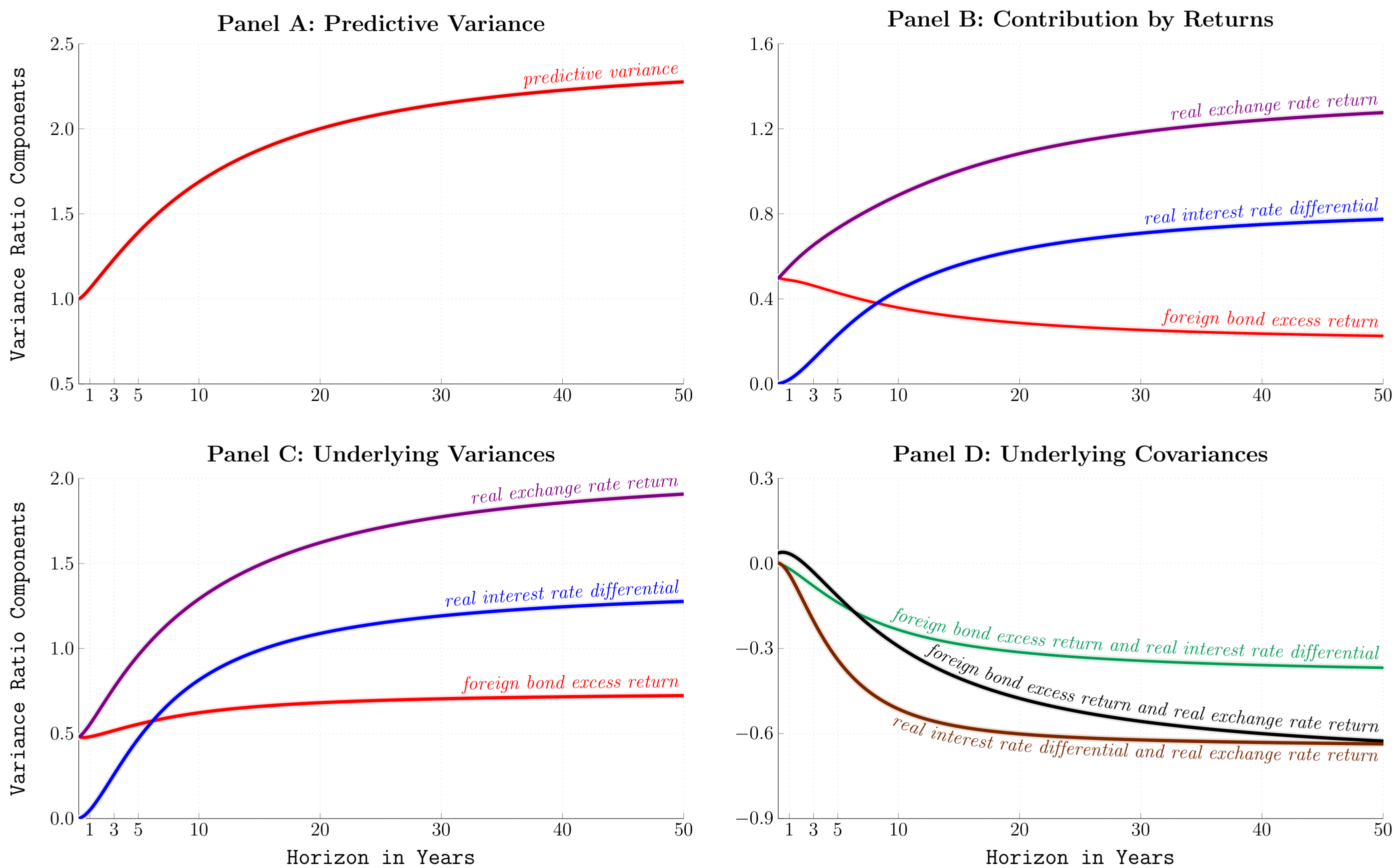

Figure 1: Predictive Variance and Return Components

This figure presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in British pounds while borrowing at the 3-month interest rate in US dollars. The excess return consists of the foreign bond excess return, real interest rate differential, and real exchange rate return. Panel A plots the predictive variance of the strategy, Panel B the contribution of each return component (i.e., the variance plus the covariances), Panel C the underlying variances, and Panel D the underlying covariances, each counted twice. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

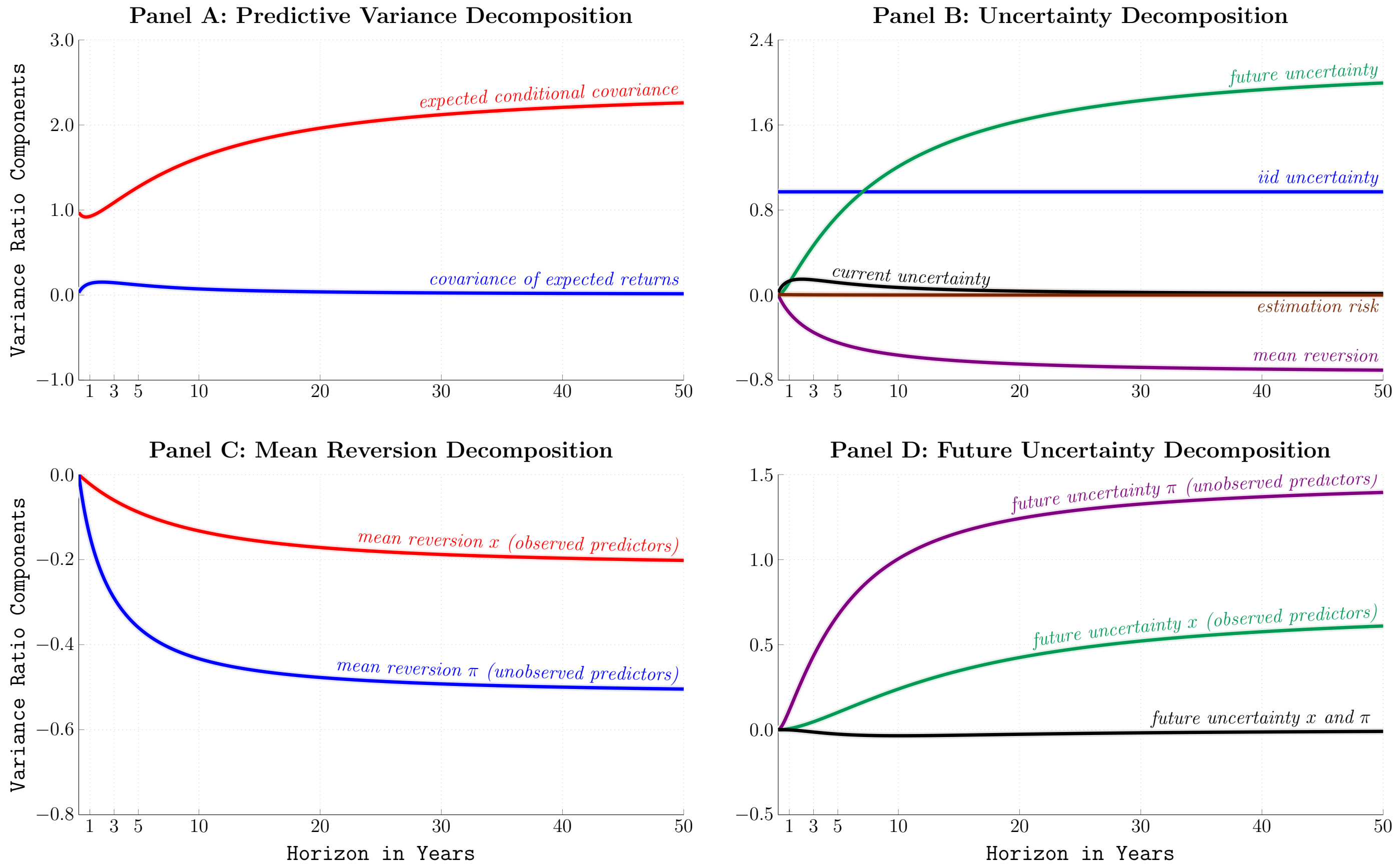

Figure 2: Predictive Variance and Uncertainty Components

This figure presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in British pounds while borrowing at the 3-month interest rate in US dollars. Panel A decomposes the predictive variance into the *expectation of the conditional covariance* and the *covariance of the conditional expectation* of future excess returns. Panel B breaks down the former into *iid uncertainty*, *mean reversion*, and *future expected return uncertainty* via Proposition 1, and the latter into *current expected return uncertainty* and *estimation risk* via Proposition 2. Panels C and D further decompose mean reversion and future uncertainty into components from observable and unobservable predictors, respectively. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

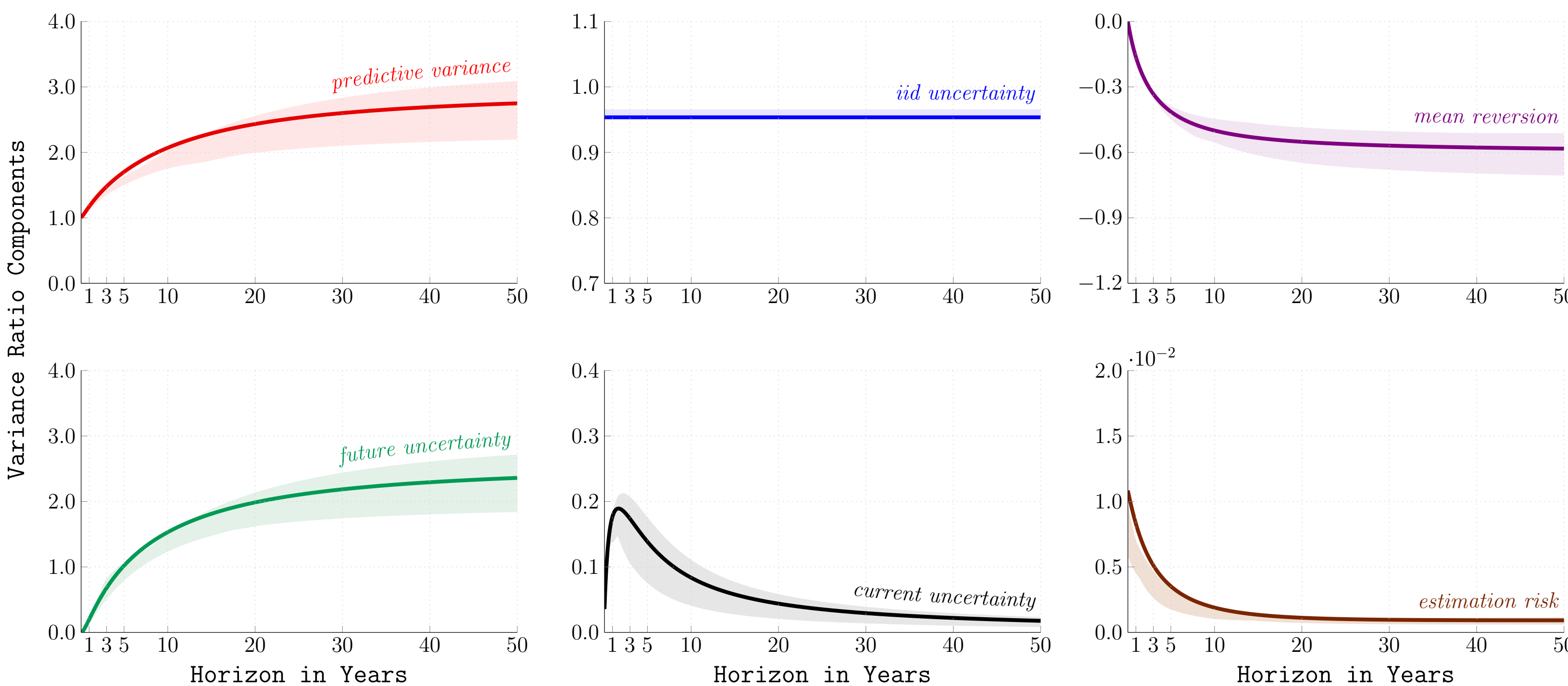

Figure 3: Predictive Variance and Uncertainty Components: A Cross-Country View

This figure presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in foreign currency while borrowing at the 3-month interest rate in US dollars. The predictive variance is decomposed into *iid uncertainty*, *mean reversion* and *future expected return uncertainty* via Proposition 1, and *current expected return uncertainty* and *estimation risk* via Proposition 2. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). The solid line denotes the cross-country average, whereas the shaded area is the interquartile range. Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, and the UK relative to the US, sourced from Global Financial Data.

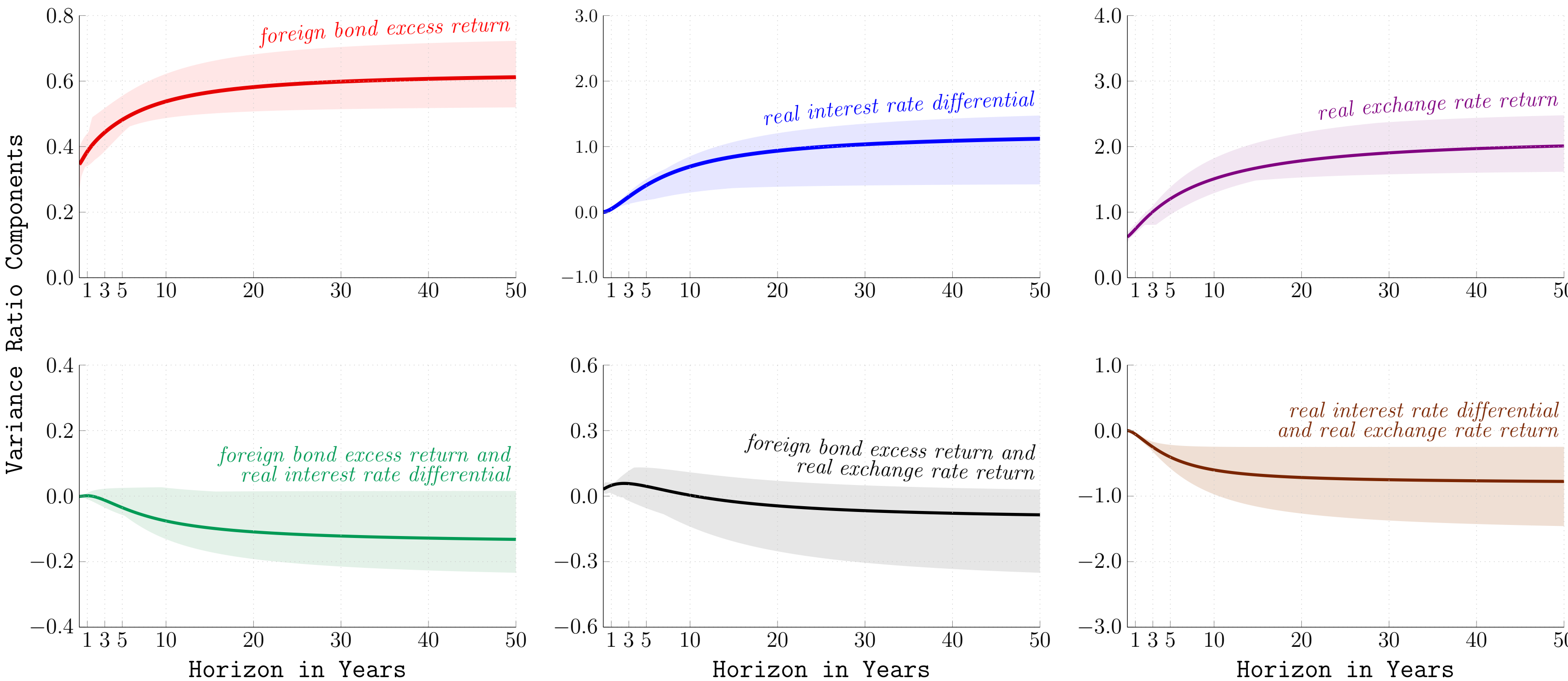

Figure 4: Predictive Variance and Return Components: A Cross-Country View

This figure presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in foreign currency while borrowing at the 3-month interest rate in US dollars. The predictive variance consists of variances and covariances (counted twice) of foreign bond excess returns, real interest rate differentials, and real exchange rate returns. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). The solid line denotes the cross-country average, whereas the shaded area is the interquartile range. Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, and the UK relative to the US, sourced from Global Financial Data.

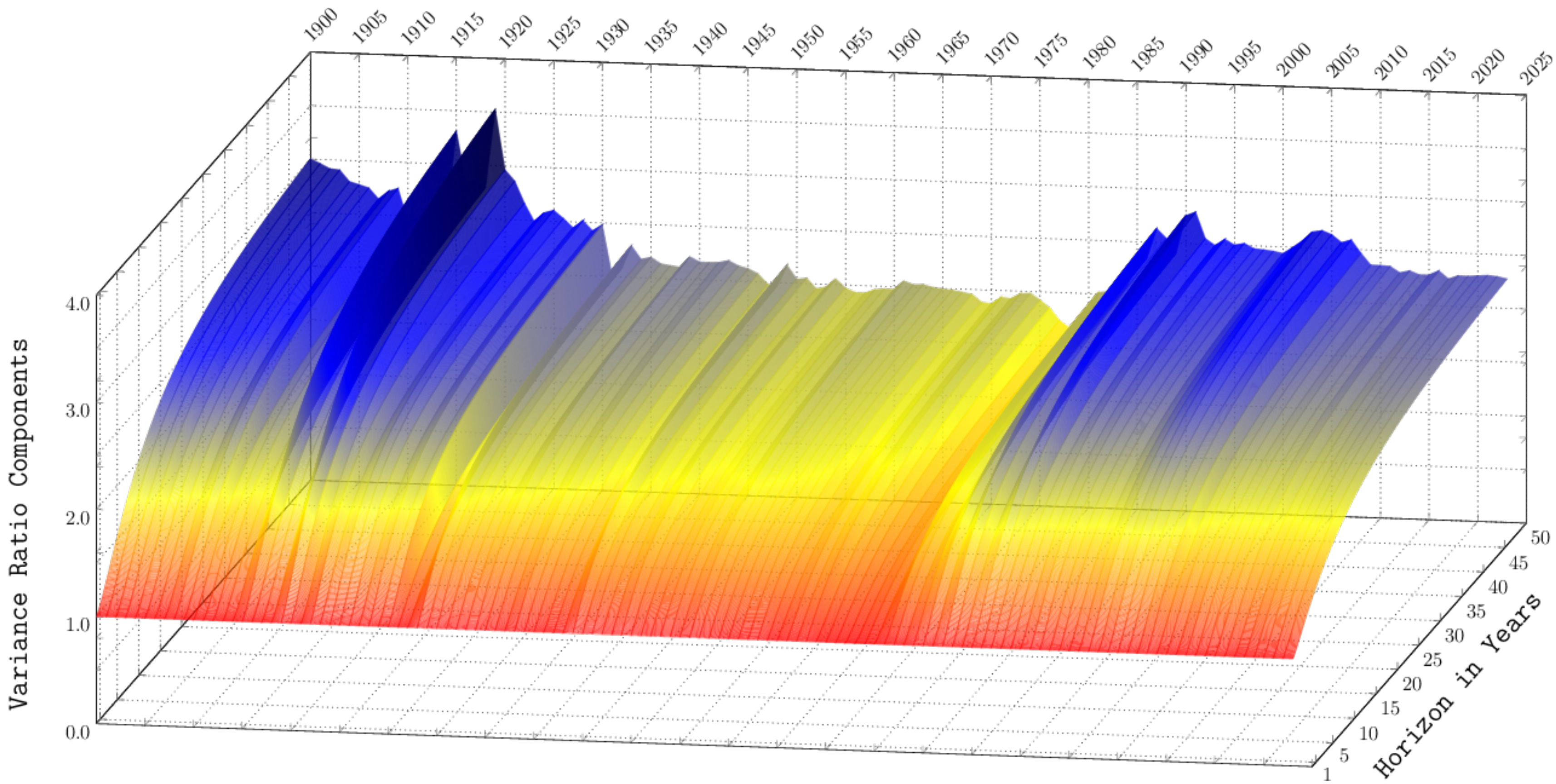

Figure 5: Predictive Variance: An Expanding Window View

This figure presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in British pounds while borrowing at the 3-month interest rate in US dollars. The analysis uses an expanding window starting in December 1900 and updating annually until December 2023. Parameter estimates are recalculated for each window using a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. For each window, we scale by the predictive variance of the one-period excess return, as described in Equation (17). The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

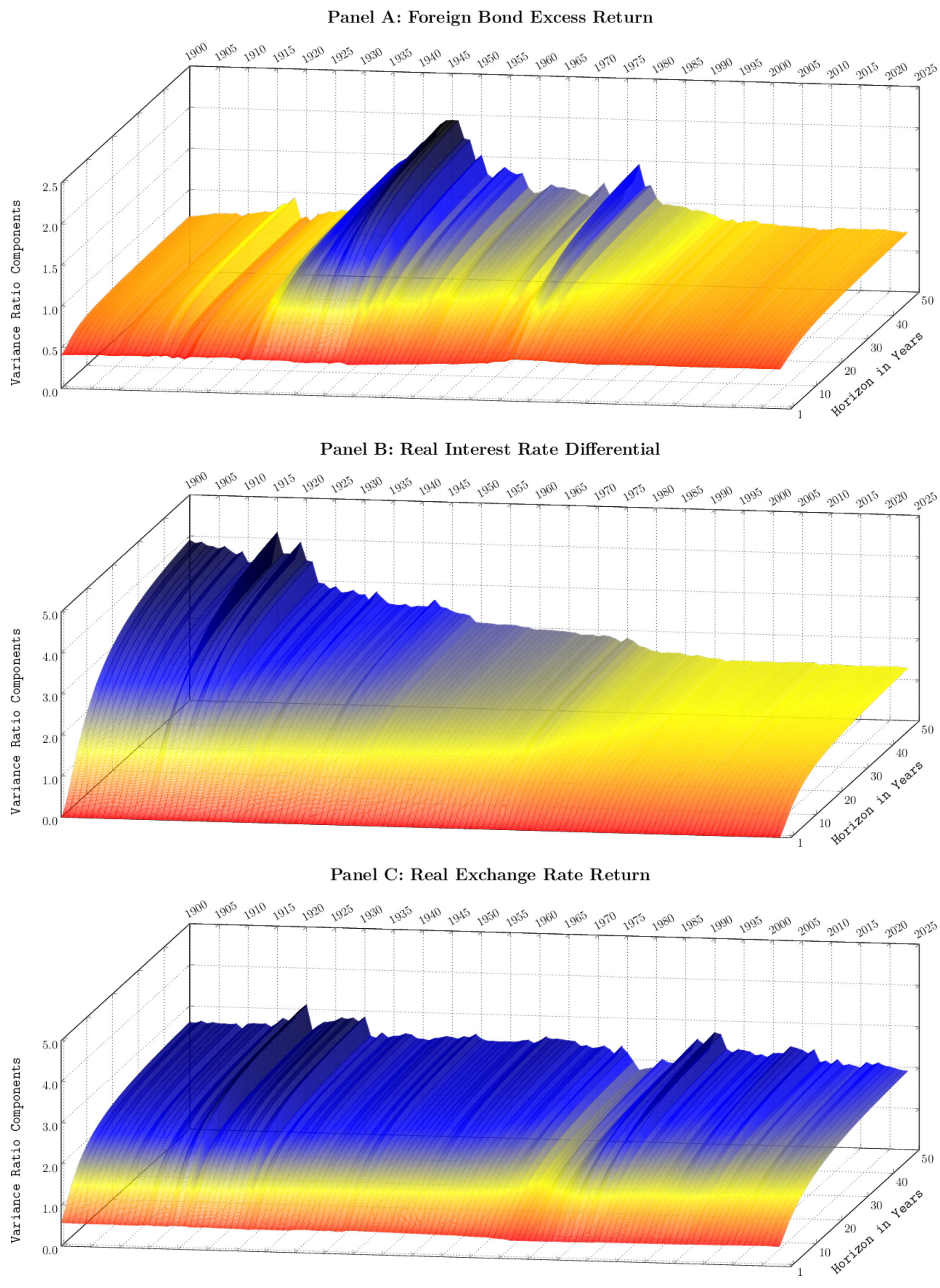

Figure 6: Return Components: An Expanding Window View

This figure presents the underlying variances contributing to the predictive variance in Figure 5. For each window, we scale each component by the predictive variance of the one-period excess return, as described in Equation (17). The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

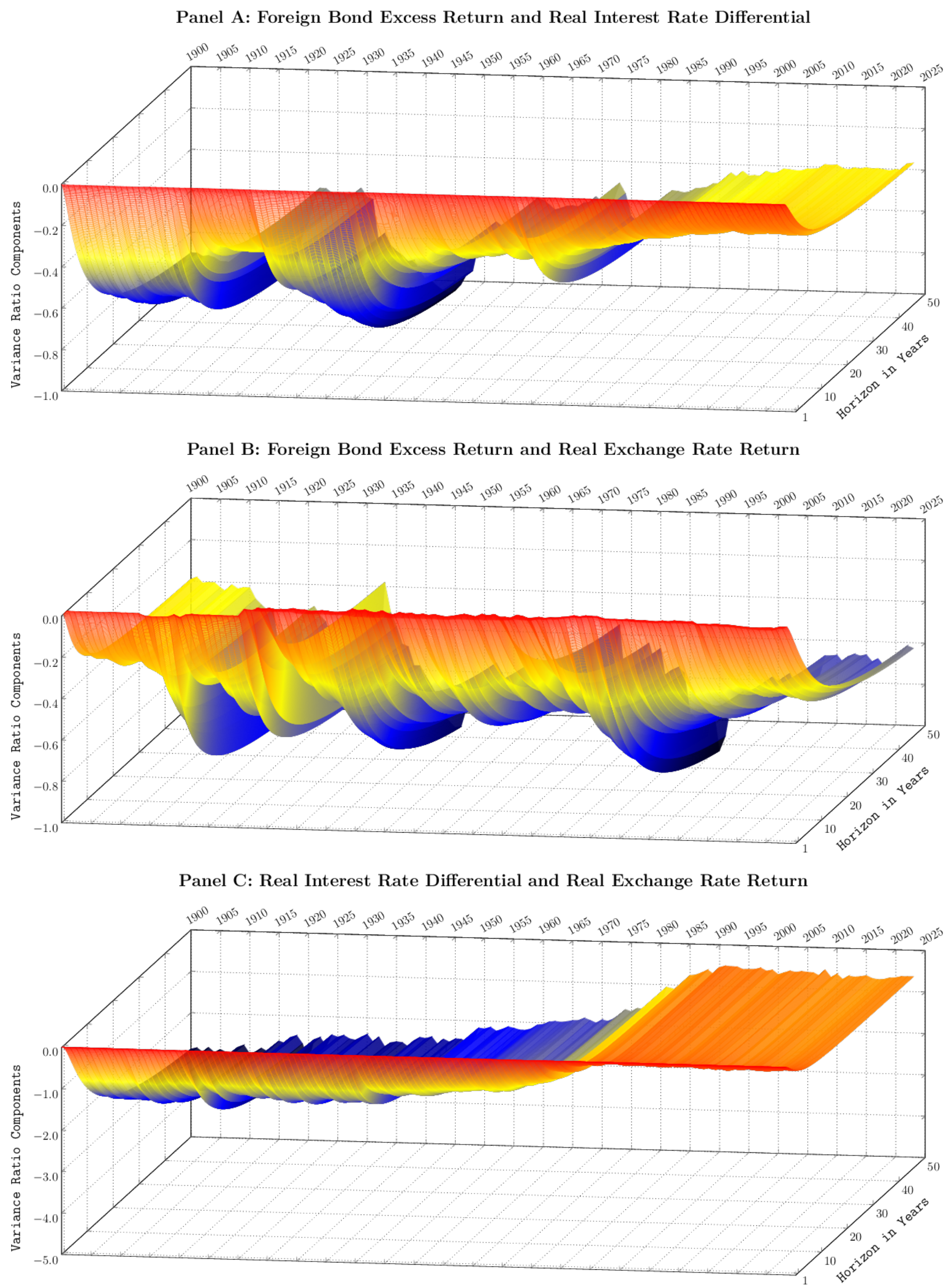

Figure 7: Return Components: An Expanding Window View

This figure presents the underlying covariances contributing to the predictive variance in Figure 5. For each window, we scale each component by the predictive variance of the one-period excess return, as described in Equation (17). The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

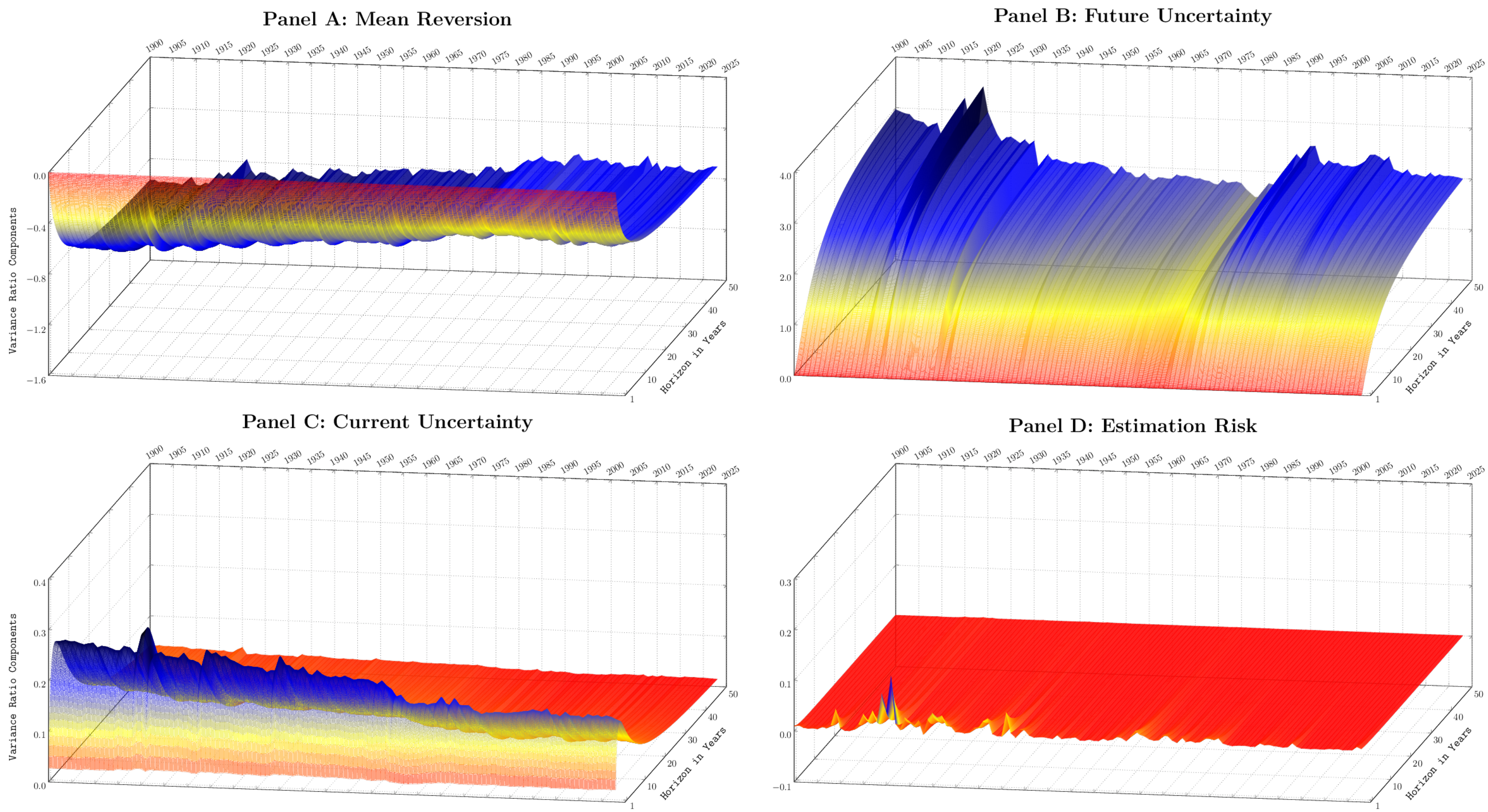

Figure 8: Uncertainty Components: An Expanding Window View

This figure presents the uncertainty components contributing to the predictive variance in Figure 5. For each window, we scale each component by the predictive variance of the one-period excess return, as described in Equation (17). The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

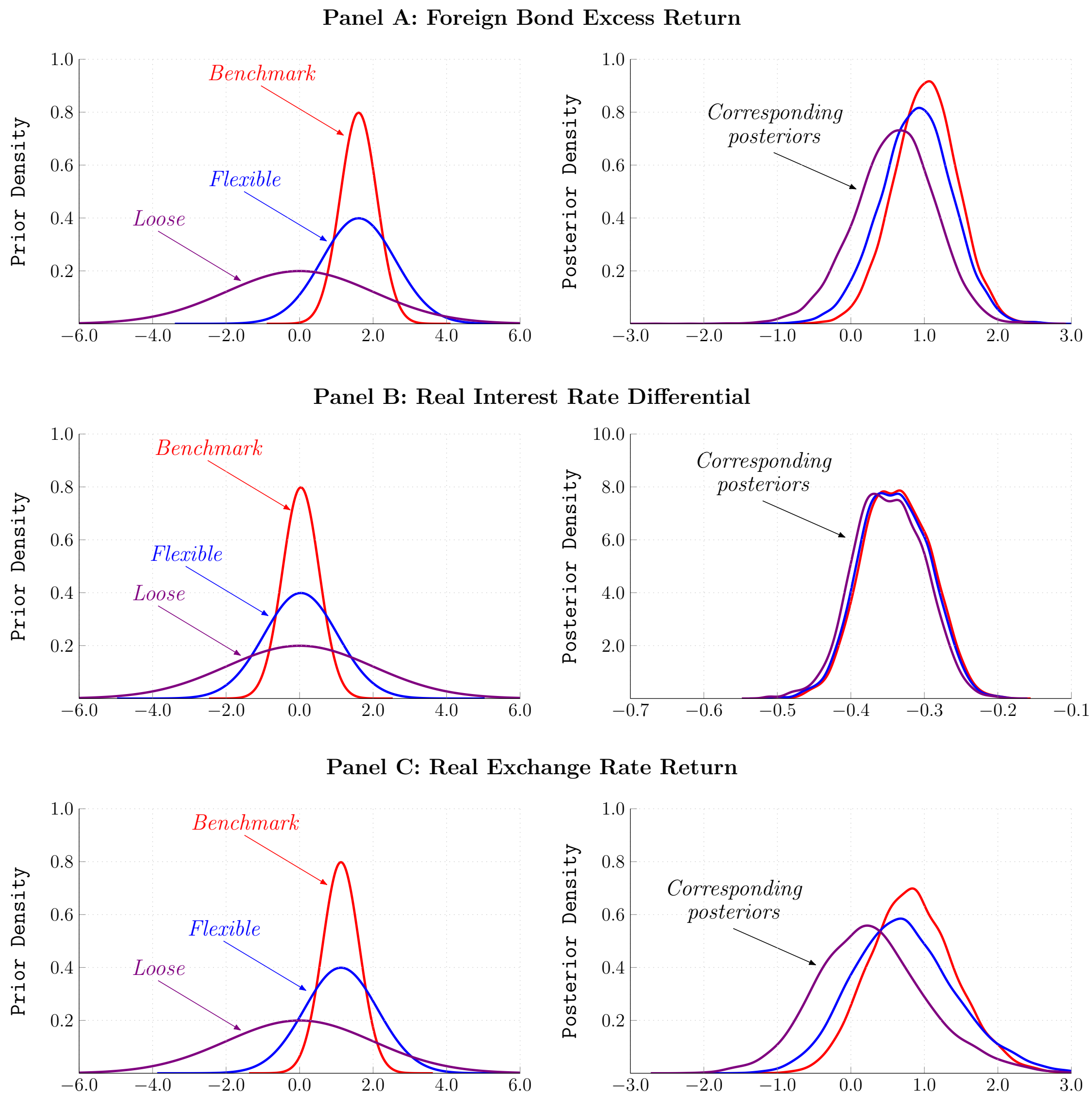

Figure 9: The Role of b Priors

This figure displays three distinct priors for the slope coefficients in b. *Benchmark* is the prior density used for the UK-US strategy, with a prior mean based on the least-squares estimate of the actual data and prior variance fixed at 0.5. *Flexible* denotes a prior density that retains the same prior mean as the *benchmark* prior but uses a prior variance of 1. *Loose* indicates a prior density with a prior mean of 0 and a prior variance of 2. Each prior follows a Normal distribution. The figure also includes the corresponding posterior densities, which are computed using a Gaussian kernel density applied to the posterior draws from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

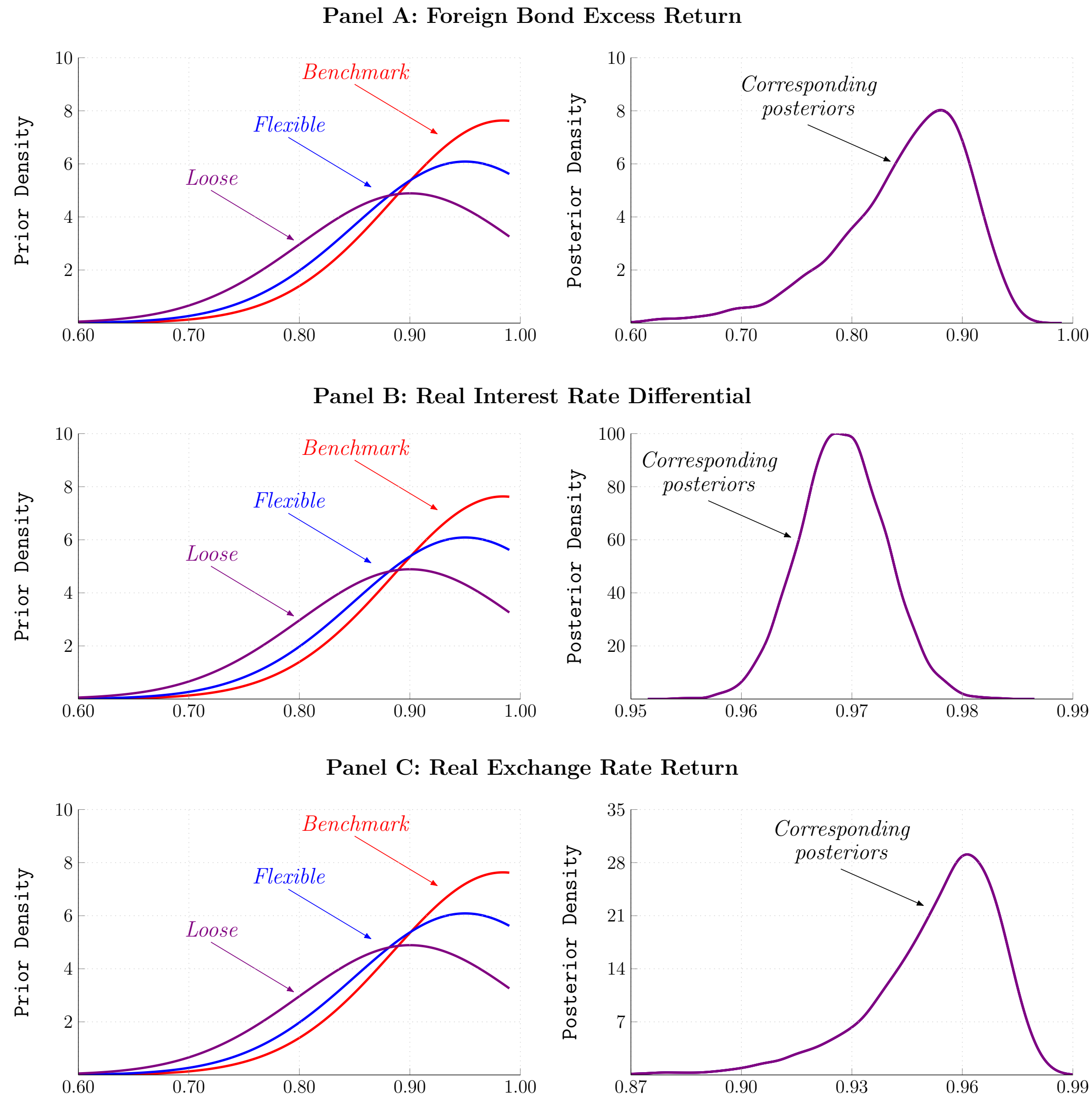

Figure 10: The Role of γ Priors

This figure displays three distinct priors for the autoregressive coefficients in γ. *Benchmark* is the prior density used for the UK-US strategy, with a prior mean based on the least-squares estimate of the actual data and prior variance fixed at 0.01. *Flexible* denotes a prior density with a prior mean of 0.95 and a prior variance of 0.01. *Loose* indicates a prior density with a prior mean of 0.90 and a prior variance of 0.01. Each prior follows a Normal distribution truncated to the interval $(-1, 1)$. The figure also includes the corresponding posterior densities, which are computed using a Gaussian kernel density applied to the posterior draws from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

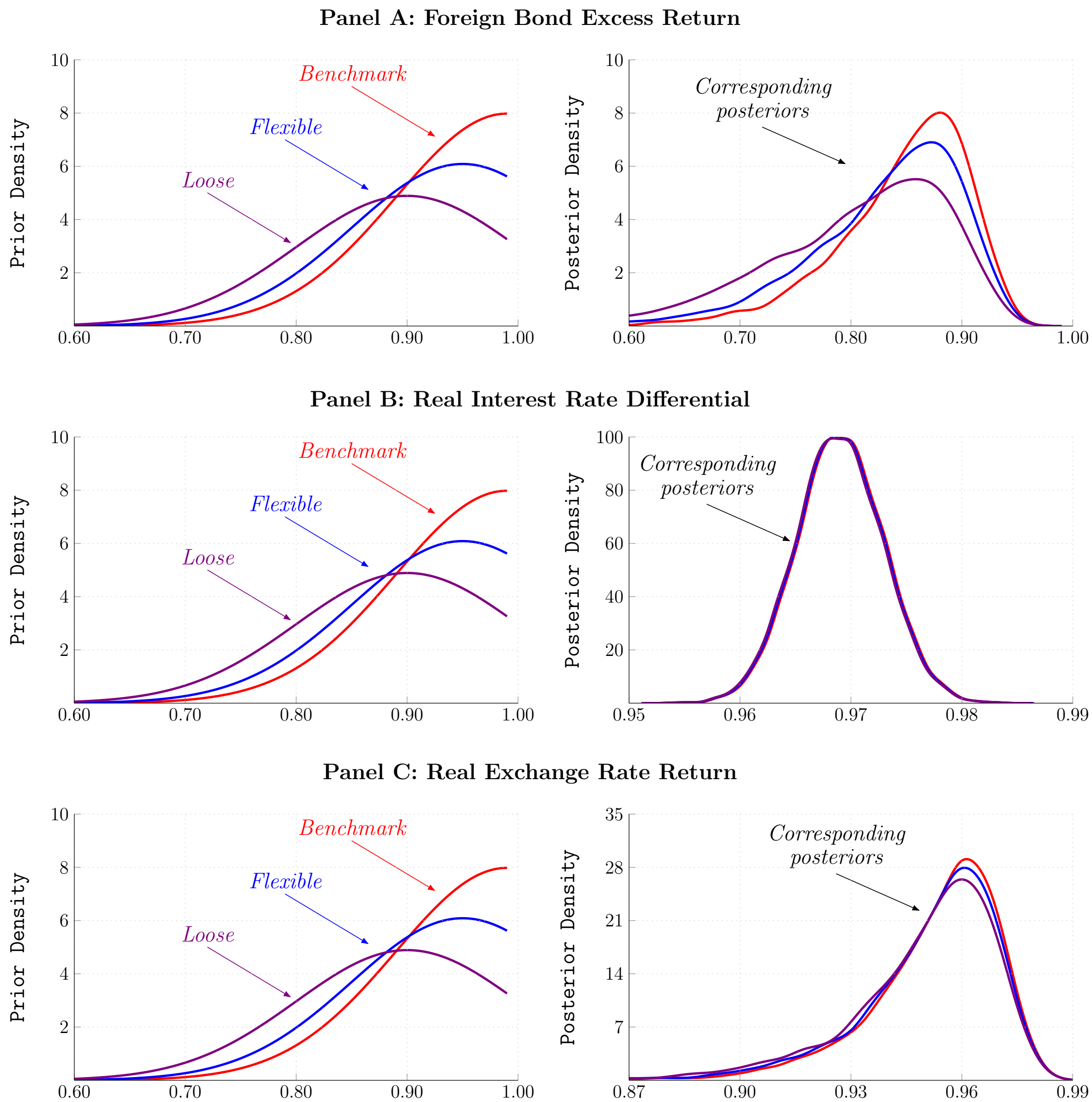

Figure 11: The Role of δ Priors

This figure displays three distinct priors for the autoregressive coefficients in δ. *Benchmark* is the prior density used for the UK-US strategy, with a prior mean of 0.99 and a prior variance fixed at 0.01. *Flexible* denotes a prior density with a prior mean of 0.95 and a prior variance of 0.01. *Loose* indicates a prior density with a prior mean of 0.90 and a prior variance of 0.01. Each prior follows a Normal distribution truncated to the interval $(0, 1)$. The figure also includes the corresponding posterior densities, which are computed using a Gaussian kernel density applied to the posterior draws from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

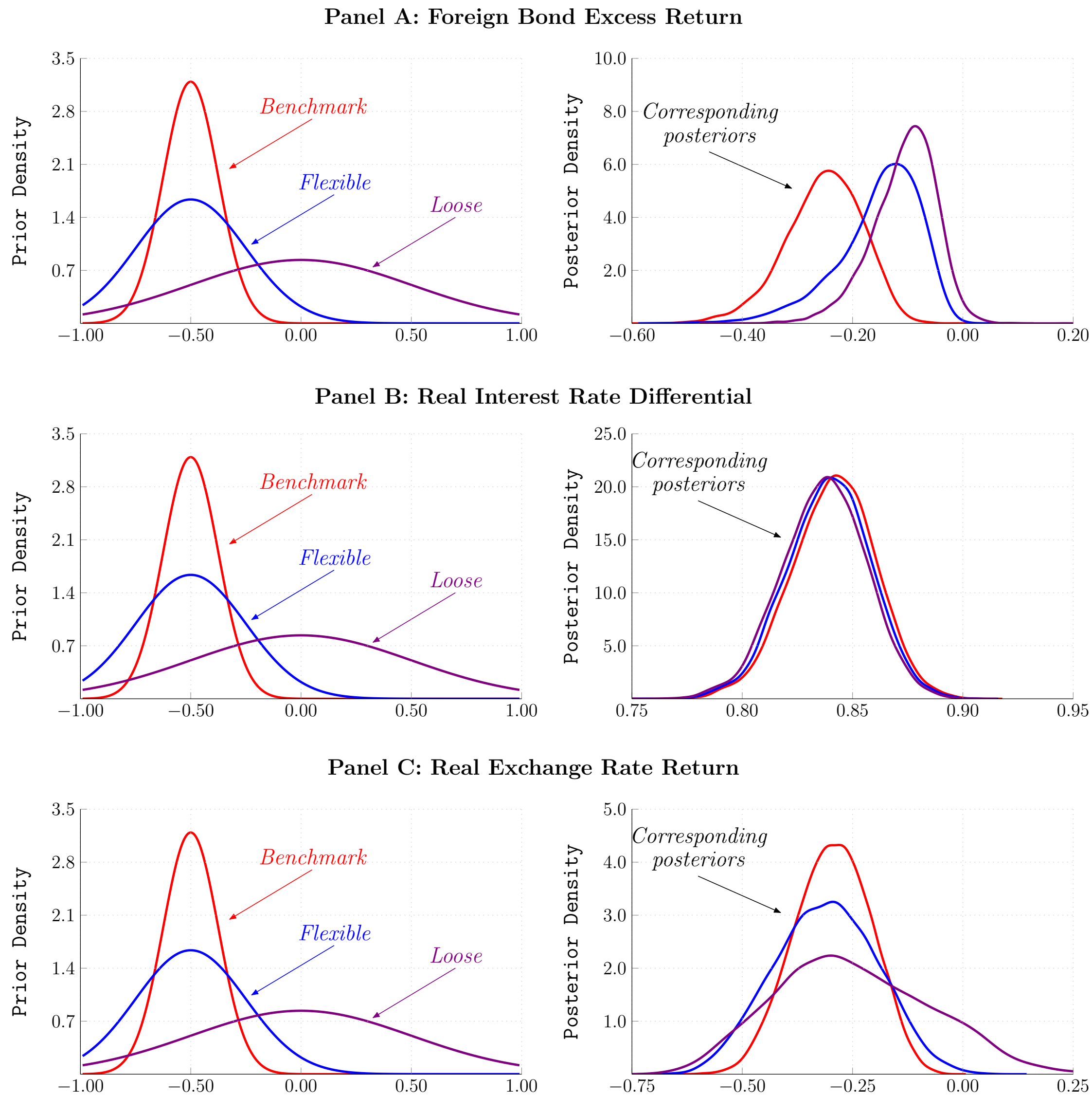

Figure 12: The Role of $\rho_{\eta u}$ Priors

This figure displays three distinct priors for the correlations between unexpected returns and shocks to the unobservable predictor in $\rho_{u\eta}$. *Benchmark* is the prior density used for the UK-US strategy, with a prior mean of -0.5 and 95% of its probability mass within the interval $[-0.75, -0.25]$. *Flexible* denotes a prior density with a prior mean of -0.50 and 95% of its probability mass within the interval $[-1, 0]$ *Loose* indicates a prior density with a prior mean of 0 and all its probability mass confined to the interval $[-1, 1]$. Each prior is derived from a Normal prior on the cross-covariance $\Sigma_{\eta u}$. The figure also includes the corresponding posterior densities, which are computed using a Gaussian kernel density applied to the posterior draws from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

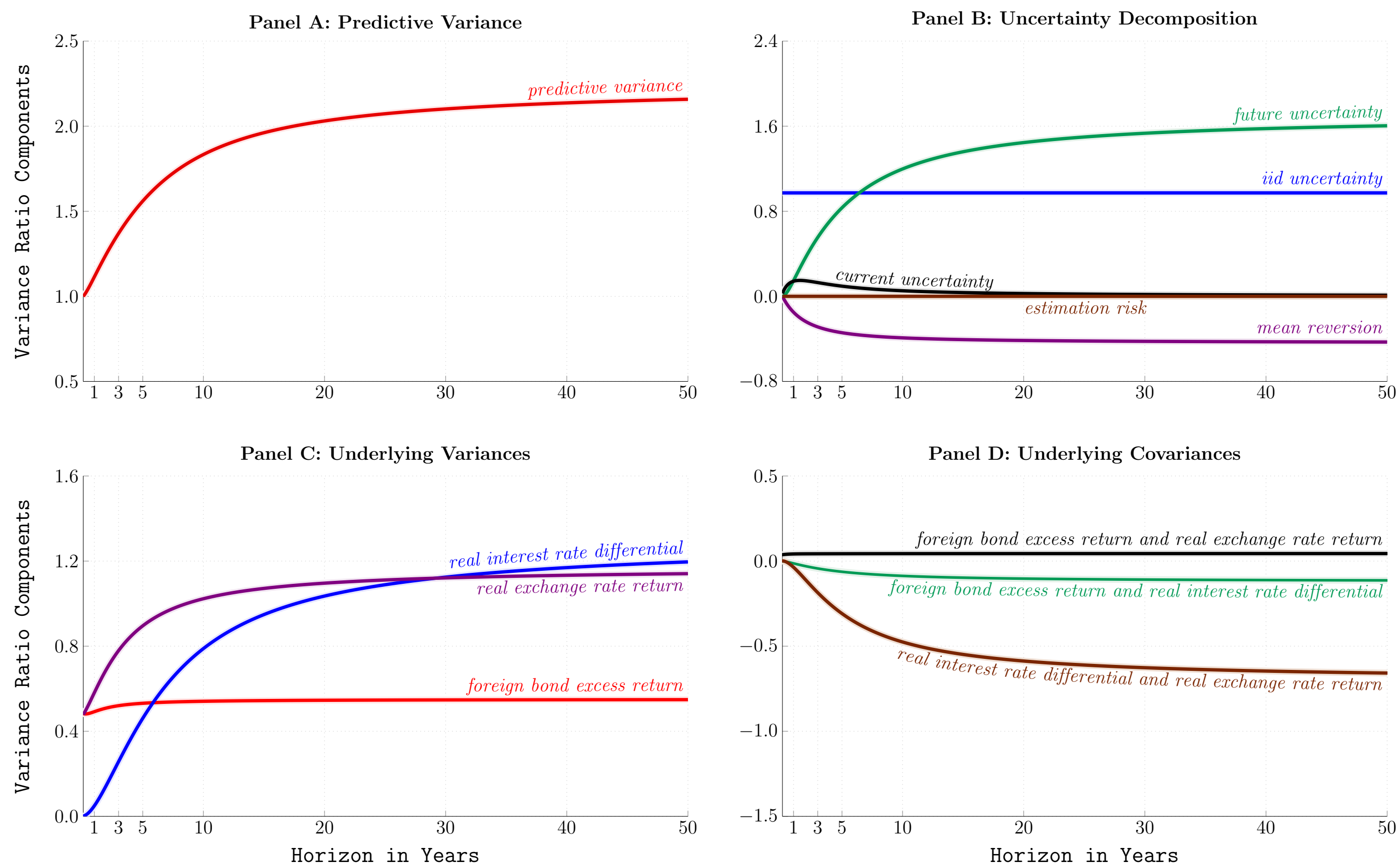

Figure 13: Predictive Variance and Unobservable Predictors

This figure presents the predictive variance based on a predictive system that only includes unobservable predictors. Panel A presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in British pounds while borrowing at the 3-month interest rate in US dollars. Panel B breaks down the predictive variance into *iid uncertainty*, *mean reversion*, and *future expected return uncertainty* via Proposition 1, and the latter into *current expected return uncertainty* and *estimation risk* via Proposition 2. Panels C and D plot the underlying variances and covariances (each counted twice) of foreign bond excess return, real interest rate differential, and real exchange rate return. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

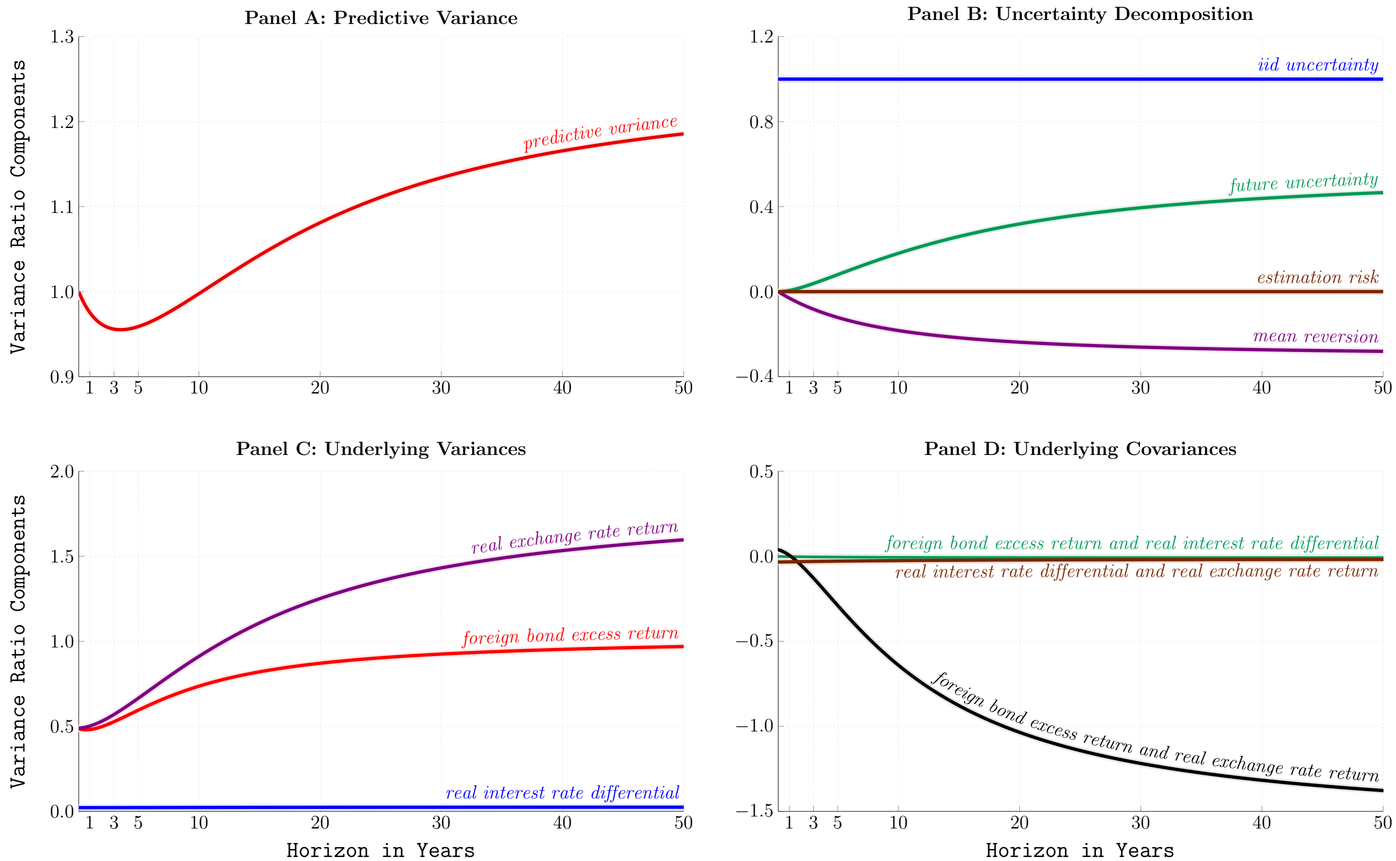

Figure 14: Predictive Variance and Observable Predictors

This figure presents the predictive variance based on a predictive system that only includes observable predictors. Panel A presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in British pounds while borrowing at the 3-month interest rate in US dollars. Panel B breaks down the predictive variance into *iid uncertainty*, *mean reversion*, and *future expected return uncertainty* via Proposition 1, and the latter into *estimation risk* via Proposition 2. Panels C and D plot the underlying variances and covariances (each counted twice) of foreign bond excess return, real interest rate differential, and real exchange rate return. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

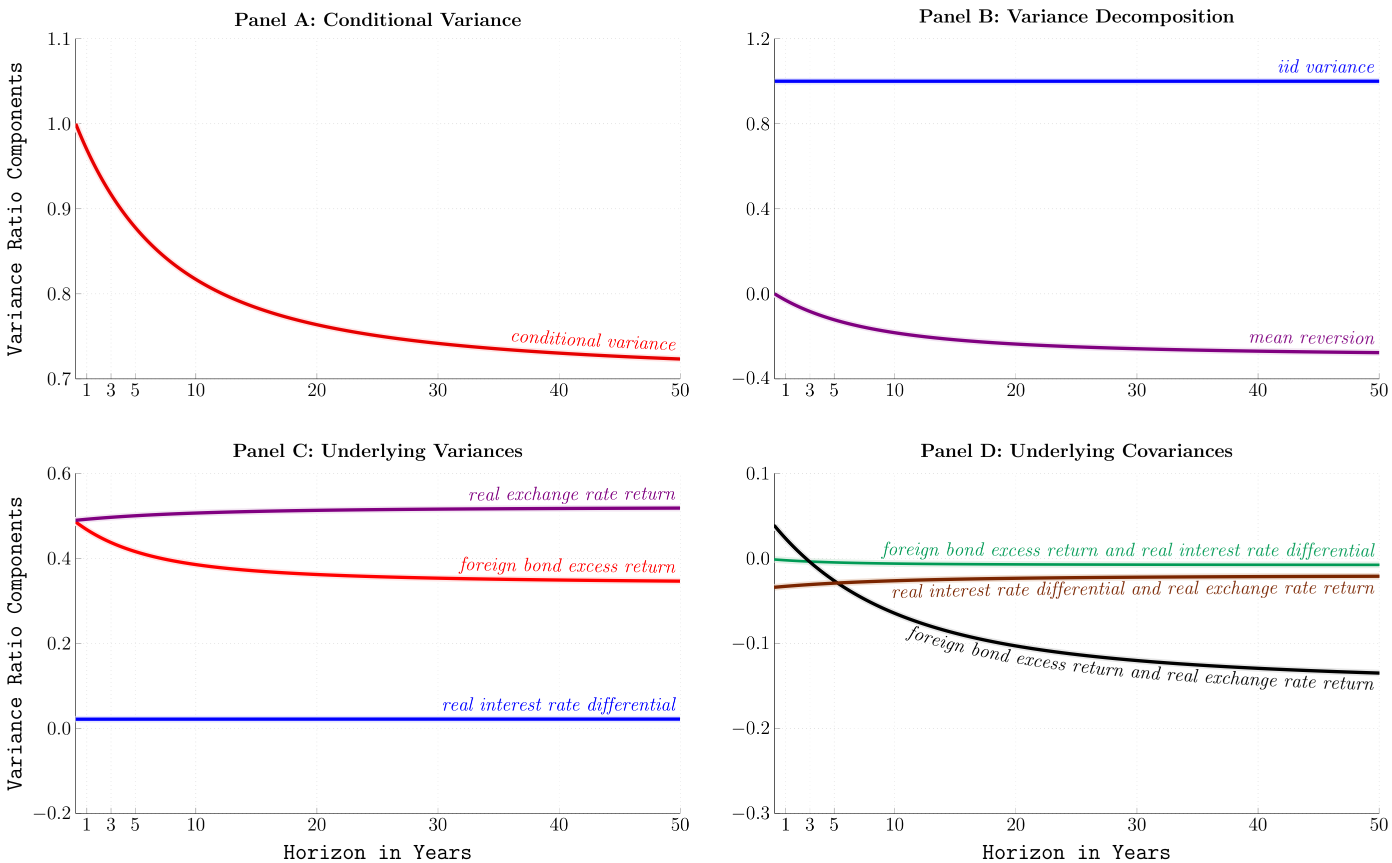

Figure 15: Conditional Variance and Observable Predictors

This figure presents the conditional variance based on a predictive system that only includes observable predictors. Panel A presents the conditional variance of the excess return for a strategy buying a 10-year constant maturity bond in British pounds while borrowing at the 3-month interest rate in US dollars. Panel B breaks down the conditional variance into *iid variance* and *mean reversion*. Panels C and D plot the underlying variances and covariances (each counted twice) of foreign bond excess return, real interest rate differential, and real exchange rate return. All components are scaled by the conditional variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

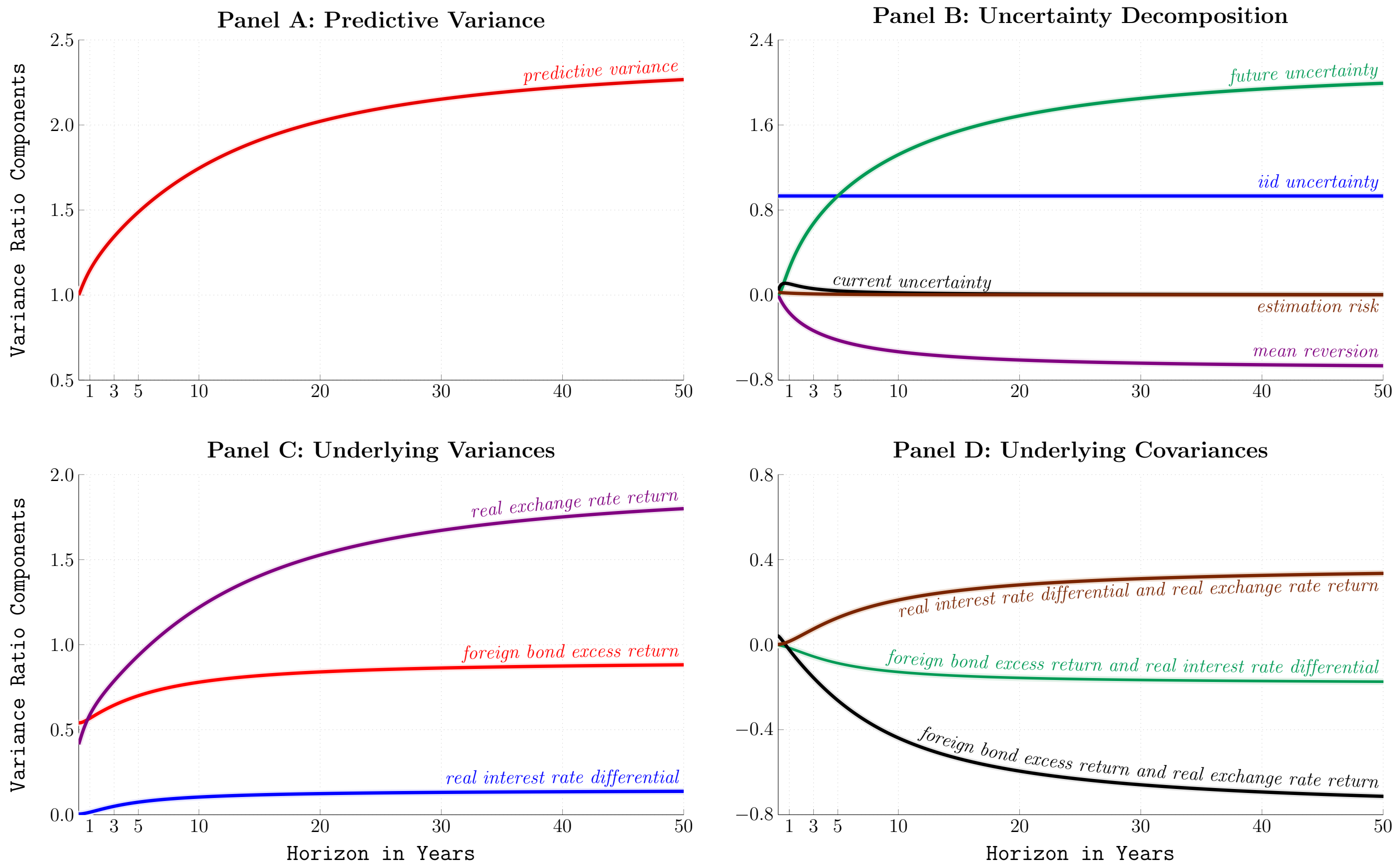

Figure 16: Predictive Variance and Return Components: Floating Regime

This figure presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in British pounds while borrowing at the 3-month interest rate in US dollars. The excess return consists of the foreign bond excess return, real interest rate differential, and real exchange rate return. Panel A plots the predictive variance of the strategy, whereas Panel B breaks down the former into *iid uncertainty*, *mean reversion*, and *future expected return uncertainty* via Proposition 1, and the latter into *current expected return uncertainty* and *estimation risk* via Proposition 2. Panels C and D plot the underlying variances and covariances (each counted twice) of foreign bond excess return, real interest rate differential, and real exchange rate return, respectively. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from January 1973 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

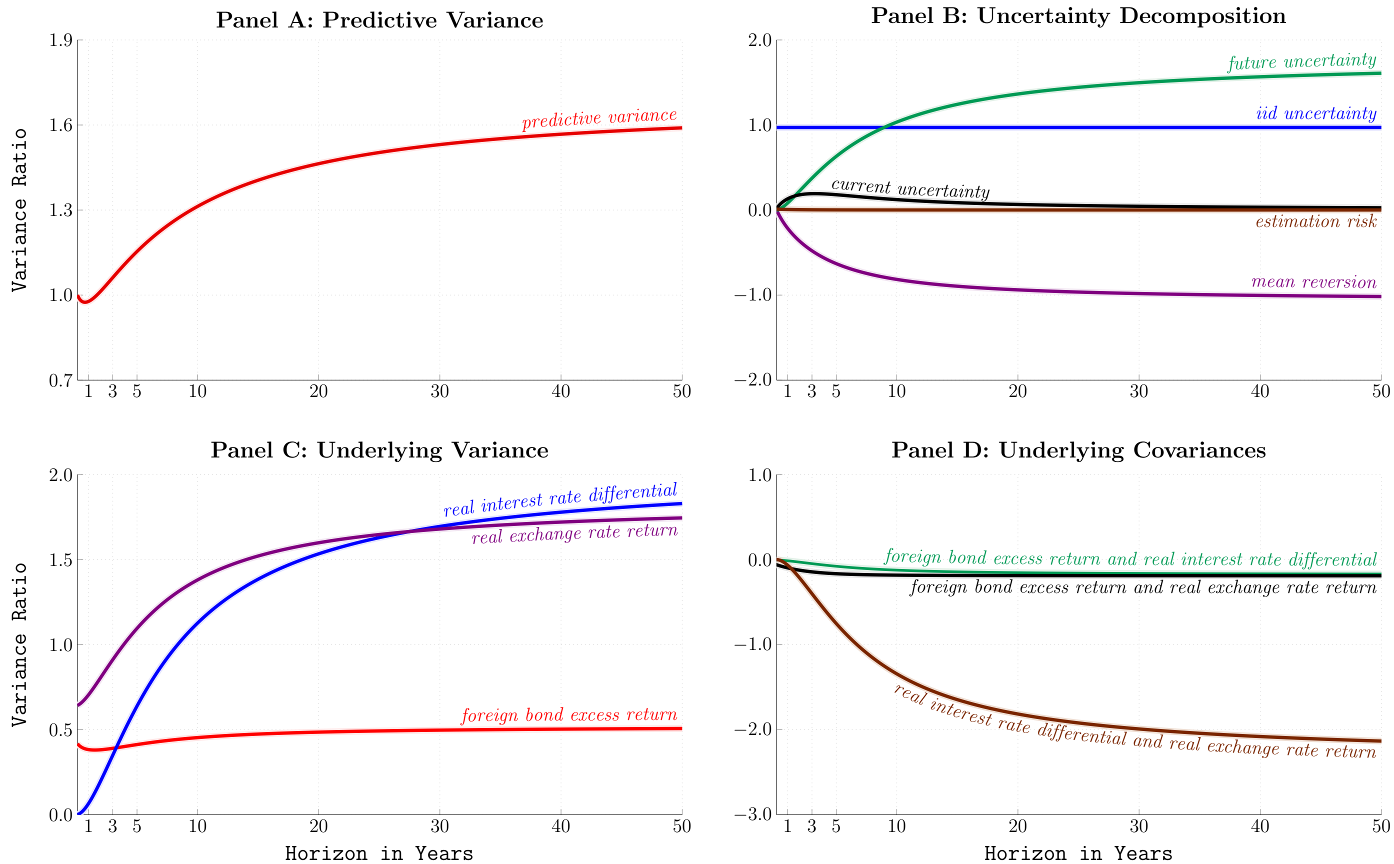

Figure 17: Predictive Variance and Return Components: Investing in US Bonds

This figure presents the predictive variance of the excess return for a strategy buying a 10-year constant maturity bond in US dollars while borrowing at the 3-month interest rate in British pounds. The excess return consists of the foreign bond excess return, real interest rate differential, and real exchange rate return. Panel A plots the predictive variance of the strategy, whereas Panel B breaks down the former into *iid uncertainty*, *mean reversion*, and *future expected return uncertainty* via Proposition 1, and the latter into *current expected return uncertainty* and *estimation risk* via Proposition 2. Panels C and D plot the underlying variances and covariances (each counted twice) of foreign bond excess return, real interest rate differential, and real exchange rate return, respectively. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the US relative to the UK, sourced from Global Financial Data.

Table 1: Summary Statistics

This table presents descriptive statistics for the dependent variables (Local Bond Excess Returns, real interest rate differentials, and real exchange rate returns) and observed predictors (Local Term Spread, real output growth differential, and nominal interest rate differential) across nine major countries relative to the US. The mean and standard deviation (*sdev*) are expressed as percentages per annum, while skewness (*skew*) and excess kurtosis (*kurt*) are robust to the impact of outliers (e.g., Kim and White, 2004). The first-order autocorrelation coefficient is denoted by ac_1. The sample period for monthly observations is indicated in parentheses, and data are sourced from Global Financial Data.

	Australia (11/1862–12/2023)					**Canada** (12/1899–12/2023)				
	mean	sdev	skew	kurt	ac_1	mean	sdev	skew	kurt	ac_1
Bond Excess Return	1.53	6.80	−0.01	4.16	−0.11	0.85	5.59	−0.03	2.44	0.05
Real Interest Rate Differential	0.29	1.47	−0.04	1.47	0.98	0.81	0.74	0.07	1.40	0.92
Real Exchange Rate Return	−0.72	16.18	−0.01	9.21	−0.27	−0.26	5.24	−0.04	4.71	−0.06
Term Spread	1.37	0.33	−0.08	1.58	0.95	0.98	0.49	−0.16	0.34	0.98
Real Output Growth Differential	−0.19	1.49	−0.05	0.95	0.99	0.45	0.92	0.04	1.08	0.98
Nominal Interest Rate Differential	0.57	0.65	0.25	0.56	0.98	0.82	0.39	0.28	0.41	0.96
	Germany (12/1814–12/2023)					**Japan** (09/1882–12/2023)				
Bond Excess Return	1.26	7.35	−0.01	1.99	0.00	1.02	7.60	0.00	3.47	−0.01
Real Interest Rate Differential	−0.43	2.15	−0.07	3.23	0.98	0.65	2.27	0.03	2.61	0.97
Real Exchange Rate Return	0.46	10.22	0.00	5.07	0.12	0.00	9.08	−0.04	4.34	0.10
Term Spread	1.29	0.41	−0.09	−0.02	0.90	0.60	0.54	−0.19	0.22	0.96
Real Output Growth Differential	−1.08	2.39	0.05	0.89	0.99	0.25	1.69	0.04	1.16	0.99
Nominal Interest Rate Differential	−0.13	0.57	0.03	−0.10	0.94	1.21	0.98	−0.17	−0.21	0.98
	New Zealand (12/1922–12/2023)					**Norway** (11/1831–12/2023)				
Bond Excess Return	−0.51	7.58	−0.01	2.83	0.08	0.30	7.98	0.03	2.16	0.01
Real Interest Rate Differential	1.69	1.05	−0.17	−0.05	0.96	0.61	1.70	0.03	1.32	0.98
Real Exchange Rate Return	−0.17	11.69	−0.03	7.05	0.02	−0.13	9.71	−0.02	5.73	0.11
Term Spread	−0.44	0.60	−0.16	−0.63	0.95	0.43	0.48	−0.07	−0.71	0.98
Real Output Growth Differential	−0.30	1.20	−0.07	1.51	0.99	−0.88	1.63	−0.01	1.88	0.99
Nominal Interest Rate Differential	2.96	0.91	0.04	1.17	0.97	0.89	0.58	−0.02	1.18	0.97
	Sweden (09/1853–12/2023)					**Switzerland** (10/1899–12/2023)				
Bond Excess Return	0.79	5.28	0.01	2.81	0.02	0.88	5.12	−0.02	2.04	0.07
Real Interest Rate Differential	0.35	1.52	0.02	1.82	0.96	0.38	1.05	0.02	1.36	0.96
Real Exchange Rate Return	−0.37	8.31	−0.03	6.71	0.06	0.73	8.32	0.04	6.24	0.06
Term Spread	0.96	0.31	0.08	1.96	0.93	0.98	0.32	−0.13	0.53	0.96
Real Output Growth Differential	−1.18	1.49	−0.05	0.89	0.99	−0.89	1.57	−0.02	1.66	0.99
Nominal Interest Rate Differential	0.61	0.52	0.31	2.55	0.97	−0.63	0.65	−0.22	0.18	0.98
	UK (12/1799 – 12/2023)									
Bond Excess Return	0.26	8.42	0.00	2.37	−0.10					
Real Interest Rate Differential	0.11	1.78	−0.05	2.17	0.98					
Real Exchange Rate Return	−0.39	8.45	−0.01	4.45	0.10					
Term Spread	0.53	0.43	0.05	0.65	0.94					
Real Output Growth Differential	−1.66	1.27	−0.03	0.85	0.99					
Nominal Interest Rate Differential	0.34	0.51	0.08	1.46	0.95					

Table 2: Posterior Estimates: Return Equations

This table presents the Bayesian posterior means, standard deviations (STD), numerical standard errors (NSE), and relative numerical inefficiency (RNI) for Equation (2). r_{t+1} includes the bond excess return in local currency, the real interest rate differential, and the real exchange rate return. x_t comprises the term spread, the real output differential, and the nominal interest rate differential. The posterior moments are obtained using a Gibbs sampling algorithm that runs for 100,000 iterations, following an initial burn-in of 20,000 iterations and retaining one in every ten iterations. The NSE provides a measure of numerical precision and reflects the variability in the posterior mean estimate if the simulation were repeated multiple times. The RNI quantifies the efficiency loss when calculating the posterior mean from autocorrelated samples, compared to independent samples. Superscripts *, **, and *** indicate that the 90%, 95%, and 99% highest posterior density intervals, respectively, exclude zero. The sample periods for each country are given in Table 1 and consists of monthly observations. Data are sourced from Global Financial Data.

		a_1	a_2	a_3	b_1	b_2	b_3	R_1^2	R_2^2	R_3^2
Australia	Mean	−0.851	0.284**	−1.109*	1.740***	0.465***	0.590	3.4	97.3	1.0
	STD	(0.640)	(0.110)	(0.643)	(0.556)	(0.044)	(0.544)			
	NSE	0.011	0.003	0.007	0.014	0.001	0.011			
	RNI	2.713	7.271	1.084	5.930	11.168	3.795			
Canada	Mean	−0.778	0.957***	−0.804	1.624***	−0.416***	0.723	8.7	92.2	5.8
	STD	(0.538)	(0.102)	(0.558)	(0.393)	(0.050)	(0.567)			
	NSE	0.006	0.003	0.008	0.006	0.002	0.014			
	RNI	1.201	6.455	1.826	2.407	9.391	5.902			
Germany	Mean	−1.141**	−0.029	0.625	1.812***	0.571***	0.641	3.4	97.6	6.1
	STD	(0.566)	(0.138)	(0.623)	(0.410)	(0.037)	(0.557)			
	NSE	0.008	0.004	0.006	0.008	0.001	0.013			
	RNI	1.958	7.395	1.012	4.196	11.411	5.651			
Japan	Mean	−0.022	0.291*	−0.586	1.697***	−0.428***	0.271	2.3	96.2	9.4
	STD	(0.523)	(0.175)	(0.656)	(0.402)	(0.052)	(0.476)			
	NSE	0.006	0.003	0.007	0.008	0.001	0.006			
	RNI	0.995	6.174	1.075	2.799	9.254	2.784			
New Zealand	Mean	0.166	1.481***	−4.320***	1.008**	−0.145***	1.342***	6.5	94.4	4.4
	STD	(0.571)	(0.113)	(0.679)	(0.456)	(0.041)	(0.376)			
	NSE	0.006	0.003	0.007	0.008	0.002	0.008			
	RNI	1.080	7.945	1.169	3.416	16.815	4.054			
Norway	Mean	−0.200	0.368***	−1.134*	1.356***	−0.535***	1.079**	4.4	97.5	4.8
	STD	(0.540)	(0.127)	(0.629)	(0.469)	(0.057)	(0.552)			
	NSE	0.006	0.004	0.007	0.009	0.002	0.013			
	RNI	1.040	7.941	1.269	3.995	11.719	5.683			
Sweden	Mean	−0.389	0.140	−0.674	1.193***	−0.195***	0.552	3.3	95.5	5.8
	STD	(0.492)	(0.160)	(0.601)	(0.454)	(0.049)	(0.558)			
	NSE	0.008	0.004	0.006	0.010	0.002	0.012			
	RNI	2.745	7.054	1.060	5.060	9.603	4.967			
Switzerland	Mean	−0.955	0.322***	1.522***	1.825***	0.095***	1.136***	14.3	94.9	6.8
	STD	(0.596)	(0.104)	(0.642)	(0.570)	(0.029)	(0.612)			
	NSE	0.007	0.003	0.007	0.009	0.001	0.014			
	RNI	1.263	5.975	1.074	2.644	8.356	5.579			
UK	Mean	−0.348	−0.216*	−0.725	1.012**	−0.337***	0.845	2.4	97.5	5.4
	STD	(0.490)	(0.128)	(0.595)	(0.434)	(0.046)	(0.582)			
	NSE	0.005	0.004	0.006	0.008	0.001	0.016			
	RNI	1.148	8.028	1.183	3.513	10.488	7.448			

Table 3: Posterior Estimates: Observed Predictor Equations

This table presents the Bayesian posterior means, standard deviations (STD), numerical standard errors (NSE), and relative numerical inefficiency (RNI) for the parameters underlying Equation (3). x_{t+1} includes the term spread, the real output differential, and the nominal interest rate differential. The posterior moments are obtained using a Gibbs Sampling algorithm that runs for 100,000 iterations, following an initial burn-in of 20,000 iterations and retaining one in every ten iterations. The NSE provides a measure of numerical precision and reflects the variability in the posterior mean estimate if the simulation were repeated multiple times. The RNI quantifies the efficiency loss when calculating the posterior mean from autocorrelated samples, compared to independent samples. Superscripts *, **, and *** indicate that the 90%, 95%, and 99% highest posterior density intervals, respectively, exclude zero. The sample periods for each country are given in Table 1 and consists of monthly observations. Data are sourced from Global Financial Data.

		θ_1	θ_2	θ_3	γ_1	γ_2	γ_3	R_1^2	R_2^2	R_3^2
Australia	Mean	0.064***	−0.005	0.013	0.952***	0.985***	0.979***	90.2	97.2	95.5
	STD	(0.012)	(0.020)	(0.011)	(0.006)	(0.003)	(0.004)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	1.404	0.945	1.027	1.408	1.069	1.645			
Canada	Mean	0.028***	0.010	0.022*	0.971***	0.984***	0.972***	95.3	96.8	91.2
	STD	(0.011)	(0.015)	(0.012)	(0.005)	(0.004)	(0.006)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	0.943	1.009	1.034	1.032	1.056	1.190			
Germany	Mean	0.052***	−0.015	−0.003	0.960***	0.987***	0.985***	80.4	97.8	87.9
	STD	(0.014)	(0.025)	(0.014)	(0.004)	(0.002)	(0.003)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	1.019	0.998	1.022	1.055	0.992	1.104			
Japan	Mean	0.022*	−0.001	0.009	0.967***	0.986***	0.986***	91.1	97.6	96.7
	STD	(0.014)	(0.022)	(0.015)	(0.005)	(0.003)	(0.003)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	0.972	0.954	1.009	1.055	1.071	1.060			
New Zealand	Mean	−0.019	0.008	0.068***	0.956***	0.980***	0.976***	90.6	97.8	94.4
	STD	(0.018)	(0.018)	(0.025)	(0.006)	(0.004)	(0.005)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	1.044	0.980	1.098	1.050	1.064	1.133			
Norway	Mean	0.008	−0.002	0.018*	0.981***	0.988***	0.978***	95.2	98.4	94.3
	STD	(0.008)	(0.015)	(0.010)	(0.003)	(0.002)	(0.004)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	1.001	0.919	1.042	1.059	1.026	1.188			
Sweden	Mean	0.058***	−0.008	0.013	0.939***	0.988***	0.979***	87.0	98.3	93.1
	STD	(0.010)	(0.015)	(0.011)	(0.006)	(0.002)	(0.004)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	0.970	0.984	1.017	1.086	0.980	1.139			
Switzerland	Mean	0.042***	−0.006	−0.011	0.958***	0.988***	0.987***	92.3	98.4	96.1
	STD	(0.010)	(0.017)	(0.012)	(0.006)	(0.001)	(0.002)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	1.181	1.012	1.012	1.235	1.018	1.008			
UK	Mean	0.013	−0.021	0.006	0.975***	0.986***	0.984***	89.0	97.3	89.5
	STD	(0.010)	(0.015)	(0.011)	(0.004)	(0.003)	(0.003)			
	NSE	0.000	0.000	0.000	0.000	0.000	0.000			
	RNI	1.051	0.945	1.027	1.150	0.934	1.550			

Table 4: Posterior Estimates: Unobserved Predictor Equations

This table presents the Bayesian posterior means, standard deviations (STD), numerical standard errors (NSE), and relative numerical inefficiency (RNI) for the parameters underlying Equation (4). The posterior moments are obtained using a Gibbs Sampling algorithm that runs for 100,000 iterations, following an initial burn-in of 20,000 iterations and retaining one in every ten iterations. The NSE provides a measure of numerical precision and reflects the variability in the posterior mean estimate if the simulation were repeated multiple times. The RNI quantifies the efficiency loss when calculating the posterior mean from autocorrelated samples, compared to independent samples. Superscripts *, **, and *** indicate that the 90%, 95%, and 99% highest posterior density intervals, respectively, exclude zero. The sample periods for each country are given in Table 1 and consists of monthly observations. Data are sourced from Global Financial Data.

		δ_1	δ_2	δ_3	R_1^2	R_2^2	R_3^2
Australia	Mean	0.914***	0.969***	0.906***	88.2	96.0	90.6
	STD	(0.048)	(0.005)	(0.057)			
	NSE	0.001	0.000	0.002			
	RNI	9.310	1.679	9.281			
Canada	Mean	0.821***	0.930***	0.906***	75.1	92.5	90.3
	STD	(0.053)	(0.008)	(0.035)			
	NSE	0.001	0.000	0.001			
	RNI	7.321	1.447	7.674			
Germany	Mean	0.892***	0.970***	0.957***	81.0	96.0	95.8
	STD	(0.041)	(0.004)	(0.010)			
	NSE	0.001	0.000	0.000			
	RNI	7.365	1.894	5.328			
Japan	Mean	0.863***	0.967***	0.948***	72.9	96.2	96.3
	STD	(0.055)	(0.006)	(0.020)			
	NSE	0.001	0.000	0.001			
	RNI	7.062	1.603	8.311			
New Zealand	Mean	0.821***	0.956***	0.914***	85.7	92.9	91.3
	STD	(0.058)	(0.008)	(0.036)			
	NSE	0.002	0.000	0.001			
	RNI	7.145	1.302	6.361			
Norway	Mean	0.888***	0.972***	0.959***	87.8	96.9	95.3
	STD	(0.043)	(0.004)	(0.013)			
	NSE	0.001	0.000	0.000			
	RNI	8.691	1.987	6.999			
Sweden	Mean	0.799***	0.961***	0.934***	72.1	96.4	94.4
	STD	(0.071)	(0.005)	(0.022)			
	NSE	0.002	0.000	0.001			
	RNI	8.815	1.345	8.445			
Switzerland	Mean	0.874***	0.947***	0.939***	89.9	93.3	94.4
	STD	(0.036)	(0.008)	(0.036)			
	NSE	0.001	0.000	0.001			
	RNI	6.940	1.404	9.768			
UK	Mean	0.847***	0.969***	0.953***	81.0	96.6	95.5
	STD	(0.059)	(0.004)	(0.018)			
	NSE	0.002	0.000	0.001			
	RNI	8.388	1.378	9.680			

Table 5: The Impact of b Priors

This table shows the impact of three distinct priors for the slope coefficients in b on the predictive variance of the UK-US strategy. *Benchmark* is the prior density used in the baseline analysis, with a prior mean based on the least-squares estimate of the actual data, and a prior variance fixed at 0.5. *Flexible* denotes a prior density that retains the same prior mean as the *benchmark* prior but uses a prior variance of 1. *Loose* indicates a prior density with a prior mean of 0 and a prior variance of 2. Each prior follows a Normal distribution. We also show the contribution of each return (i.e., the variance plus the covariances) and uncertainty components. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

Horizon in Years	1	10	20	30	50
			Benchmark		
Predictive Variance	1.06	1.69	2.00	2.15	2.28
Foreign Bond Excess Return	0.49	0.36	0.29	0.25	0.23
Real Interest Rate Differential	0.02	0.44	0.63	0.71	0.77
Real Exchange Rate Return	0.55	0.89	1.08	1.18	1.28
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.17	−0.57	−0.65	−0.68	−0.71
Future Uncertainty	0.12	1.21	1.64	1.83	2.00
Current Uncertainty	0.13	0.07	0.04	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00
			Flexible		
Predictive Variance	1.07	1.73	2.05	2.20	2.33
Foreign Bond Excess Return	0.49	0.38	0.33	0.30	0.28
Real Interest Rate Differential	0.02	0.45	0.64	0.72	0.79
Real Exchange Rate Return	0.56	0.90	1.08	1.18	1.26
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.16	−0.54	−0.61	−0.64	−0.67
Future Uncertainty	0.13	1.23	1.66	1.85	2.01
Current Uncertainty	0.14	0.07	0.04	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00
			Loose		
Predictive Variance	1.10	1.79	2.07	2.19	2.30
Foreign Bond Excess Return	0.50	0.45	0.43	0.42	0.41
Real Interest Rate Differential	0.02	0.48	0.67	0.74	0.80
Real Exchange Rate Return	0.58	0.86	0.97	1.03	1.08
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.15	−0.46	−0.51	−0.53	−0.55
Future Uncertainty	0.14	1.22	1.58	1.74	1.86
Current Uncertainty	0.14	0.06	0.03	0.02	0.01
Estimation Risk	0.01	0.00	0.00	0.00	0.00

Table 6: The Impact of γ Priors

This table shows the impact of three distinct priors for the autoregressive coefficients in γ on the predictive variance of the UK-US strategy. *Benchmark* is the prior density used in the baseline analysis, with a prior mean based on the least-squares estimate of the actual data, and a prior variance of 0.01. *Flexible* denotes a prior density with a prior mean of 0.95 and a prior variance of 0.01. *Loose* indicates a prior density with a prior mean of 0.90 and a prior variance of 0.01. Each prior follows a Normal distribution truncated to the interval $(-1, 1)$. We also show the contribution of each return (i.e., the variance plus the covariances) and uncertainty components. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

Horizon in Years	1	10	20	30	50
			Benchmark		
Predictive Variance	1.06	1.69	2.00	2.15	2.28
Foreign Bond Excess Return	0.49	0.36	0.29	0.25	0.23
Real Interest Rate Differential	0.02	0.44	0.63	0.71	0.77
Real Exchange Rate Return	0.55	0.89	1.08	1.18	1.28
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.17	−0.57	−0.65	−0.68	−0.71
Future Uncertainty	0.12	1.21	1.64	1.83	2.00
Current Uncertainty	0.13	0.07	0.04	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00
			Flexible		
Predictive Variance	1.06	1.69	2.00	2.15	2.28
Foreign Bond Excess Return	0.49	0.36	0.29	0.25	0.23
Real Interest Rate Differential	0.02	0.44	0.63	0.71	0.77
Real Exchange Rate Return	0.55	0.89	1.08	1.18	1.28
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.17	−0.57	−0.65	−0.68	−0.71
Future Uncertainty	0.12	1.21	1.64	1.83	2.00
Current Uncertainty	0.13	0.07	0.04	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00
			Loose		
Predictive Variance	1.06	1.69	2.00	2.15	2.27
Foreign Bond Excess Return	0.49	0.36	0.29	0.25	0.23
Real Interest Rate Differential	0.02	0.44	0.63	0.71	0.77
Real Exchange Rate Return	0.55	0.89	1.08	1.18	1.27
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.17	−0.57	−0.65	−0.68	−0.71
Future Uncertainty	0.12	1.21	1.64	1.83	1.99
Current Uncertainty	0.13	0.07	0.04	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00

Table 7: The Impact of δ Priors

This table shows the impact of three distinct priors for the autoregressive coefficients in δ on the predictive variance of the UK-US strategy. *Benchmark* is the prior density used in the baseline analysis, with a prior mean of 0.99 and a prior variance of 0.01. *Flexible* denotes a prior density with a prior mean of 0.95 and a prior variance of 0.01. *Loose* indicates a prior density with a prior mean of 0.90 and a prior variance of 0.01. Each prior follows a Normal distribution truncated to the interval $(0, 1)$. We also show the contribution of each return (i.e., the variance plus the covariances) and uncertainty components. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

Horizon in Years	1	10	20	30	50
			Benchmark		
Predictive Variance	1.06	1.69	2.00	2.15	2.28
Foreign Bond Excess Return	0.49	0.36	0.29	0.25	0.23
Real Interest Rate Differential	0.02	0.44	0.63	0.71	0.77
Real Exchange Rate Return	0.55	0.89	1.08	1.18	1.28
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.17	−0.57	−0.65	−0.68	−0.71
Future Uncertainty	0.12	1.21	1.64	1.83	2.00
Current Uncertainty	0.13	0.07	0.04	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00
			Flexible		
Predictive Variance	1.06	1.67	1.98	2.13	2.25
Foreign Bond Excess Return	0.49	0.35	0.28	0.24	0.22
Real Interest Rate Differential	0.02	0.44	0.63	0.71	0.78
Real Exchange Rate Return	0.55	0.88	1.07	1.17	1.26
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.16	−0.55	−0.63	−0.66	−0.69
Future Uncertainty	0.12	1.18	1.60	1.79	1.95
Current Uncertainty	0.13	0.07	0.04	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00
			Loose		
Predictive Variance	1.06	1.65	1.95	2.10	2.22
Foreign Bond Excess Return	0.49	0.34	0.27	0.23	0.20
Real Interest Rate Differential	0.02	0.45	0.64	0.71	0.78
Real Exchange Rate Return	0.55	0.87	1.05	1.15	1.24
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.16	−0.53	−0.61	−0.63	−0.66
Future Uncertainty	0.12	1.15	1.55	1.74	1.90
Current Uncertainty	0.12	0.06	0.03	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00

Table 8: The Impact of $\rho_{\eta u}$ Priors

This table shows the impact of three distinct priors for the correlations between unexpected returns and shocks to the unobservable predictor in $\rho_{u\eta}$ on the predictive variance of the UK-US strategy. *Benchmark* is the prior density used for the UK-US strategy, with a prior mean of -0.5 and 95% of its probability mass within the interval $[-0.75, -0.25]$. *Flexible* denotes a prior density with a prior mean of -0.50 and 95% of its probability mass within the interval $[-1, 0]$ *Loose* indicates a prior density with a prior mean of 0 and all its probability mass confined to the interval $[-1, 1]$. Each prior is derived from a Normal prior on the cross-covariance $\Sigma_{\eta u}$. Each prior is derived from a Normal prior on the cross-covariance $\Sigma_{\eta u}$. We also show the contribution of each return (i.e., the variance plus the covariances) and uncertainty components. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data.

Horizon in Years	1	10	20	30	50
			Benchmark		
Predictive Variance	1.06	1.69	2.00	2.15	2.28
Foreign Bond Excess Return	0.49	0.36	0.29	0.25	0.23
Real Interest Rate Differential	0.02	0.44	0.63	0.71	0.77
Real Exchange Rate Return	0.55	0.89	1.08	1.18	1.28
IID Uncertainty	0.97	0.97	0.97	0.97	0.97
Mean Reversion	−0.17	−0.57	−0.65	−0.68	−0.71
Future Uncertainty	0.12	1.21	1.64	1.83	2.00
Current Uncertainty	0.13	0.07	0.04	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00
			Flexible		
Predictive Variance	1.04	1.59	1.90	2.04	2.18
Foreign Bond Excess Return	0.47	0.28	0.19	0.14	0.11
Real Interest Rate Differential	0.02	0.46	0.66	0.74	0.81
Real Exchange Rate Return	0.55	0.85	1.05	1.16	1.26
IID Uncertainty	0.96	0.96	0.96	0.96	0.96
Mean Reversion	−0.15	−0.53	−0.61	−0.64	−0.66
Future Uncertainty	0.12	1.09	1.51	1.70	1.87
Current Uncertainty	0.12	0.06	0.03	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00
			Loose		
Predictive Variance	1.05	1.63	1.94	2.09	2.22
Foreign Bond Excess Return	0.46	0.27	0.18	0.14	0.10
Real Interest Rate Differential	0.02	0.49	0.71	0.80	0.87
Real Exchange Rate Return	0.57	0.87	1.05	1.15	1.25
IID Uncertainty	0.93	0.93	0.93	0.93	0.93
Mean Reversion	−0.13	−0.44	−0.51	−0.54	−0.56
Future Uncertainty	0.13	1.09	1.49	1.68	1.84
Current Uncertainty	0.11	0.05	0.03	0.02	0.01
Estimation Risk	0.00	0.00	0.00	0.00	0.00

Internet Appendix to

What 200 Years of Data Tell Us About the Predictive Variance of Long-Term Bonds?

(not for publication)

Abstract

We present the derivations of the Proposition 1 and Proposition 2, and a description of the Gibbs sampling algorithm.

A Decomposition of $Cov\left(r^k_{i,T}, r^k_{j,T} \mid D_T\right)$

In Equation (8), we demonstrate that the long-horizon predictive covariance can be decomposed into two parts: the expectation of the conditional covariance of k-period returns and the covariance of the conditional expectation of k-period returns. Proposition 1 reveals that the first component can be further decomposed into three primary sources of uncertainty, namely, *iid uncertainty*, *mean reversion*, and *future expected return uncertainty*. Meanwhile, Proposition 2 shows that the second component is driven by two main sources of uncertainty, i.e., : current expected return uncertainty and estimation risk. Here, we derive the expressions for these components in closed-form.

A.1 Proof of Proposition 1

Proof. Since $x_{i,t}$ (the observable predictor of return i at time t) in Equation (3) follows a first-order autoregression with $-1 < \gamma_i < 1$, we can rewrite $x_{i,t}$ as

$$x_{i,t} = \frac{1}{b_i}(E_{r_i} - a_i) + \sum_{l=0}^{\infty} \gamma_i^l v_{i,t-l}, \tag{A.1}$$

whenever $b_i \neq 0$. Similarly, since $0 < \delta_i < 1$, we can rewrite $\pi_{i,t}$ (the unobserved predictor of asset i at time t) in Equation (4) as

$$\pi_{i,t} = \sum_{l=0}^{\infty} \delta_i^l \eta_{i,t-l}. \tag{A.2}$$

From Equations (2)–(5), the k-period return of asset i can be written as

$$\begin{aligned} r_{i,T+k} =& a_i + (1-\gamma_i^{k-1})(E_{r_i} - a_i) + b_i \gamma_i^{k-1} x_{i,T} + b_i \sum_{l=1}^{k-1} \gamma_i^{k-l-1} v_{i,T+l} \\ &+ \delta_i^{k-1} \pi_{i,T} + \sum_{l=1}^{k-1} \delta_i^{k-l-1} \eta_{i,T+l} + u_{i,T+k}. \end{aligned} \tag{A.3}$$

The k-period return from period $T+1$ through period $T+k$ is then

$$r_{i,T}^{k} = \sum_{l=1}^{k} r_{i,T+l} = k\mathbb{E}_{r_i} + \frac{1-\gamma_i^k}{1-\gamma_i}(a_i + b_i x_{i,T} - \mathbb{E}_{r_i}) + b_i \sum_{l=1}^{k-1} \frac{1-\gamma_i^{k-l}}{1-\gamma_i} v_{i,T+l}$$
$$+ \frac{1-\delta_i^k}{1-\delta_i}\pi_{i,T} + \sum_{l=1}^{k-1} \frac{1-\delta_i^{k-l}}{1-\delta_i}\eta_{i,T+l} + \sum_{l=1}^{k} u_{i,T+l}. \tag{A.4}$$

The conditional covariance $\mathbb{C}\mathrm{ov}\left(r_{i,T}^{k}, r_{j,T}^{k} \mid \pi_T, \phi, D_T\right)$ of the k-period returns $r_{i,T}^{k}$ and $r_{j,T}^{k}$ can be obtained from Equation (A.4) as

$$\begin{aligned}
&\mathbb{C}\mathrm{ov}\left(r_{i,T}^{k}, r_{j,T}^{k} \mid \pi_T, \phi, D_T\right) = k\sigma_{u_i u_j} + \\
&b_i b_j \frac{\sigma_{v_i v_j}}{(1-\gamma_i)(1-\gamma_j)} \left(k - 1 - \gamma_i \frac{1-\gamma_i^{k-1}}{1-\gamma_i} - \gamma_j \frac{1-\gamma_j^{k-1}}{1-\gamma_j} + \gamma_i\gamma_j \frac{1-\gamma_i^{k-1}\gamma_j^{k-1}}{1-\gamma_i\gamma_j}\right) + \\
&b_i \frac{\sigma_{v_i \eta_j}}{(1-\gamma_i)(1-\delta_j)} \left(k - 1 - \gamma_i \frac{1-\gamma_i^{k-1}}{1-\gamma_i} - \delta_j \frac{1-\delta_j^{k-1}}{1-\delta_j} + \gamma_i\delta_j \frac{1-\gamma_i^{k-1}\delta_j^{k-1}}{1-\gamma_i\delta_j}\right) + \\
&b_j \frac{\sigma_{\eta_i v_j}}{(1-\delta_i)(1-\gamma_j)} \left(k - 1 - \delta_i \frac{1-\delta_i^{k-1}}{1-\delta_i} - \gamma_j \frac{1-\gamma_j^{k-1}}{1-\gamma_j} + \delta_i\gamma_j \frac{1-\delta_i^{k-1}\gamma_j^{k-1}}{1-\delta_i\gamma_j}\right) + \\
&\frac{\sigma_{\eta_i \eta_j}}{(1-\delta_i)(1-\delta_j)} \left(k - 1 - \delta_i \frac{1-\delta_i^{k-1}}{1-\delta_i} - \delta_j \frac{1-\delta_j^{k-1}}{1-\delta_j} + \delta_i\delta_j \frac{1-\delta_i^{k-1}\delta_j^{k-1}}{1-\delta_i\delta_j}\right) + \\
&b_i \frac{\sigma_{v_i u_j}}{1-\gamma_i} \left(k - 1 - \gamma_i \frac{1-\gamma_i^{k-1}}{1-\gamma_i}\right) + b_j \frac{\sigma_{u_i v_j}}{1-\gamma_j} \left(k - 1 - \gamma_j \frac{1-\gamma_j^{k-1}}{1-\gamma_j}\right) + \\
&\frac{\sigma_{\eta_i u_j}}{1-\delta_i} \left(k - 1 - \delta_i \frac{1-\delta_i^{k-1}}{1-\delta_i}\right) + \frac{\sigma_{u_i \eta_j}}{1-\delta_j} \left(k - 1 - \delta_j \frac{1-\delta_j^{k-1}}{1-\delta_j}\right).
\end{aligned} \tag{A.5}$$

We can simplify Equation (A.5) by setting

$$A_\chi(k) = 1 + \frac{1}{k}\left(-1 - \chi\frac{1-\chi^{k-1}}{1-\chi}\right)$$
$$B_{\chi\psi}(k) = 1 + \frac{1}{k}\left(-1 - \chi\frac{1-\chi^{k-1}}{1-\chi} - \psi\frac{1-\psi^{k-1}}{1-\psi} + \chi\psi\frac{1-\chi^{k-1}\psi^{k-1}}{1-\chi\psi}\right)$$

for $\chi = \{\gamma_i, \gamma_j, \delta_i, \delta_j\}$ and $\chi\psi = \{\gamma_i\gamma_j, \delta_i\delta_j, \gamma_i\delta_j, \delta_i\gamma_j\}$, and obtaining

$$\bar{d}_i = \left(\frac{1+\delta_i}{1-\delta_i}\frac{R_i^2}{1-R_i^2}\frac{\sigma_{\pi_i}^2}{\sigma_{r_i}^2 - \sigma_{u_i}^2}\right)^{1/2}$$

$$\bar{e}_i = \left(\frac{1+\gamma_i}{1-\gamma_i}\frac{R_i^2}{1-R_i^2}\frac{\sigma_{x_i}^2}{\sigma_{r_i}^2 - \sigma_{u_i}^2}\right)^{1/2}$$

for $s = \{i, j\}$ from the following relationships

$$\sigma_{\eta_i}^2 = \sigma_{\pi_i}^2(1-\delta_i^2) = \sigma_{r_i}^2(1-\delta_i^2)R_i^2\frac{\sigma_{\pi_i}^2}{\sigma_{r_i}^2 - \sigma_{u_i}^2} = \sigma_{u_i}^2(1-\delta_i^2)\frac{R_i^2}{1-R_i^2}\frac{\sigma_{\pi_i}^2}{\sigma_{r_i}^2 - \sigma_{u_i}^2},$$

$$\sigma_{v_i}^2 = \sigma_{x_i}^2(1-\gamma_i^2) = \sigma_{r_i}^2(1-\gamma_i^2)R_i^2\frac{\sigma_{x_i}^2}{\sigma_{r_i}^2 - \sigma_{u_i}^2} = \sigma_{u_i}^2(1-\gamma_i^2)\frac{R_i^2}{1-R_i^2}\frac{\sigma_{x_i}^2}{\sigma_{r_i}^2 - \sigma_{u_i}^2},$$

which hold when returns are predictable. Finally, take the expectation with respect to D_T and break up the expectation of the conditional covariance into the three sources given by Equations (10)–(12). □

A.2 Proof of Proposition 2

Proof. We finally derive a closed-form expression for the second term of the right side of Equation (8). For ease of exposition, let $\mathbb{E}_{i,T}^k = \mathbb{E}\left(r_{i,T}^k \mid \pi_T, \phi, D_T\right)$ and $\mathbb{E}_{j,T}^k = \mathbb{E}\left(r_{j,T}^k \mid \pi_T, \phi, D_T\right)$. The covariance of $\mathbb{E}_{i,T}^k$ and $\mathbb{E}_{j,T}^k$ given D_T can be decomposed as

$$\begin{aligned}\mathbb{C}\text{ov}\left(\mathbb{E}_{i,T}^k, \mathbb{E}_{j,T}^k \mid D_T\right) = \mathbb{E}\left[Cov\left(\mathbb{E}_{i,T}^k, \mathbb{E}_{j,T}^k \mid \phi, D_T\right) \mid D_T\right] \\ + \mathbb{C}\text{ov}\left[\mathbb{E}\left(\mathbb{E}_{i,T}^k \mid \phi, D_T\right), \mathbb{E}\left(\mathbb{E}_{j,T}^k \mid \phi, D_T\right) \mid D_T\right]. \qquad \text{(A.6)}\end{aligned}$$

Analogous to Pástor and Stambaugh (2012), let $c_{i,T}$ be the conditional mean of the unobserved predictor $\pi_{i,T}$ on ϕ and D_T and denote as $q_{ij,T}$ the conditional covariance between the unobserved predictors $\pi_{i,T}$ and $\pi_{j,T}$ as

$$c_{i,T} = E(\pi_{i,T} \mid \phi, D_T)$$
$$q_{ij,T} = Cov(\pi_{i,T},\, \pi_{j,T} \mid \phi, D_T).$$

By Equation (A.4),

$$\mathbb{E}^k_{i,T} = k\mathbb{E}_{r_i} + \frac{1-\gamma_i^k}{1-\gamma_i}(a_i + b_i x_{i,T} - \mathbb{E}_{r_i}) + \frac{1-\delta_i^k}{1-\delta_i}\pi_{i,T}.$$

It follows that

$$\mathbb{E}\left(\mathbb{E}^k_{i,T} \mid \phi, D_T\right) = k\mathbb{E}_{r_i} + \frac{1-\gamma_i^k}{1-\gamma_i}(a_i + b_i x_{i,T} - \mathbb{E}_{r_i}) + \frac{1-\delta_i^k}{1-\delta_i}c_{i,T}, \tag{A.7}$$

and

$$\mathbb{C}\mathrm{ov}\left(\mathbb{E}^k_{i,T}, \mathbb{E}^k_{j,T} \mid \phi, D_T\right) = \frac{1-\delta_i^k}{1-\delta_i}\frac{1-\delta_j^k}{1-\delta_j}q_{ij,T}. \tag{A.8}$$

Substituting Equation (A.7) and Equation (A.8) into Equation (A.6) and adding the conditional expectation on D_T gives Equation (13). Note that $\mathbb{E}_{r_i} = a_i + b_i\theta_i/(1-\gamma_i)$ and $b_{i,T} = a_i + b_i x_{i,T} - \mathbb{E}_{r_s}$.

$\square$

B Bayesian Estimation

B.1 Predictive System in Compact Form

The predictive system in Equations (2)–(5) can equivalently be written as

$$y_t = X_{t-1}\beta + Z\pi_{t-1} + \varepsilon_t \tag{A.9}$$

$$\pi_t = \delta\pi_{t-1} + \eta_t, \tag{A.10}$$

with

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \overset{iid}{\sim} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{\varepsilon\varepsilon} & \Sigma'_{\eta\varepsilon} \\ \Sigma_{\eta\varepsilon} & \Sigma_{\eta\eta} \end{bmatrix}\right). \tag{A.11}$$

The vector y_t, which combines returns and observed predictors, is given by

$$y_t = \begin{bmatrix} r_t \\ x_t \end{bmatrix}_{2m\times 1}$$

with m denoting of number of returns.

The matrix X_{t-1}, which includes the lagged observable predictors, is defined as

$$X_{t-1} = \begin{bmatrix} X_{1,t-1} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & X_{m,t-1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & X_{1,t-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & X_{m,t-1} \end{bmatrix}_{2m\times 4m}$$

with $X_{i,t-1} = [1, x_{i,t-1}]$.

The vector β is defined as

$$\beta = \begin{bmatrix} \text{vec}\begin{pmatrix} a' \\ \text{diag}(b)' \end{pmatrix} \\ \\ \text{vec}\begin{pmatrix} \theta' \\ \text{diag}(\gamma)' \end{pmatrix} \end{bmatrix}_{4m\times 1},$$

where diag(b) denotes the diagonal elements of b, the matrix of slope coefficients with zero off-diagonal elements, whereas diag(γ) refers to the diagonal elements of γ, the matrix of first-order autoregressive coefficients with zero off-diagonal elements.

The vector ε_t collects unexpected returns and shocks to observable predictors as follows

$$\varepsilon_t = \begin{bmatrix} u_t \\ v_t \end{bmatrix}_{2m\times 1}$$

and its covariance matrix $\Sigma_{\varepsilon\varepsilon}$ is defined as

$$\Sigma_{\varepsilon\varepsilon} = \begin{bmatrix} \Sigma_{uu} & \Sigma'_{vu} \\ \Sigma_{vu} & \Sigma_{vv} \end{bmatrix}_{2m\times 2m}. \tag{A.12}$$

Finally, the matrix Z is given by

$$Z = \begin{bmatrix} I_m \\ 0_m \end{bmatrix}_{2m\times m},$$

where I_m is an m-dimensional identity matrix and 0_m is a m-dimensional zero matrix. In our empirical analysis, $m = 3$.

B.2 An Equivalent Representation with Orthogonal Shocks

As we want to impose a negative prior on the covariance between u_t and η_t as in Pástor and Stambaugh (2012), it is convenient to rewrite the system in Equations (A.9)–(A.11) with orthogonal shocks to y_t and π_t.

Define a zero-mean random vector ζ_t that is orthogonal to ε_t as

$$\zeta_t = \eta_t - \Sigma_{\eta\varepsilon}\Sigma_{\varepsilon\varepsilon}^{-1}\varepsilon_t$$

with covariance matrix given by

$$\Sigma_{\zeta\zeta} = \Sigma_{\eta\eta} - \Sigma_{\eta\varepsilon}\Sigma_{\varepsilon\varepsilon}^{-1}\Sigma_{\eta\varepsilon}',$$

and substitute into Equation (A.10) to obtain

$$\pi_t = \delta\pi_{t-1} + \Sigma_{\eta\varepsilon}\Sigma_{\varepsilon\varepsilon}^{-1}\varepsilon_t + \zeta_t.$$

We can then rewrite the predictive system in equations (A.9)–(A.11) as

$$y_t = X_{t-1}\beta + Z\pi_{t-1} + \varepsilon_t \tag{A.13}$$

$$\pi_t = N_{t-1}\varphi + \zeta_t \tag{A.14}$$

with

$$\begin{bmatrix} \varepsilon_t \\ \zeta_t \end{bmatrix} \overset{i.i.d.}{\sim} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{\varepsilon\varepsilon} & 0 \\ 0 & \Sigma_{\zeta\zeta} \end{bmatrix} \right). \tag{A.15}$$

The matrix N_{t-1} is defined as

$$N_t = \begin{bmatrix} N_{1,t-1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & N_{m,t} \end{bmatrix}_{m \times m(2m+1)}$$

where $N_{i,t-1} = [\pi_{i,t-1}, (\Sigma_{\varepsilon\varepsilon}^{-1}\varepsilon_t)']$, and the vector φ given by

$$\varphi = \begin{bmatrix} \text{vec}\begin{pmatrix} \text{diag}(\delta)' \\ \Sigma_{\eta\varepsilon}' \end{pmatrix} \end{bmatrix}_{m(2m+1)\times 1}$$

B.3 Summary of the Algorithm

We estimate $\phi = [\beta, \varphi, \Sigma_{\varepsilon\varepsilon}, \Sigma_{\zeta\zeta}]$, the unknown parameters of the predictive system in Equations (A.13)–(A.15), using a Gibbs Sampling algorithm that conditions on $D_T = \{y_1, \ldots, y_T\}$ and $\pi = [\pi_1, \ldots, \pi_T]'$. The latter is sampled in one block using the *forward filtering, backward sampling* approach of Carter and Kohn (1994). We draw 100,000 iterations (beyond a burn-in sample of 20,000 iterations) from the conditional posterior distributions and then keep a draw every 10 draws for a total of 10,000 draws.

The algorithm consists of the following steps:

1. Initialize $\beta, \Sigma_{\varepsilon\varepsilon}, \Sigma_{\zeta\zeta}$, and π_0 using random draws from the prior distributions,
2. Sample $\varphi \mid D_T, \pi, \phi_{-\varphi}$ from a conditional Normal posterior distribution,
3. Sample $\Sigma_{\zeta\zeta}^{-1} \mid D_T, \pi, \phi_{-\Sigma_{\zeta\zeta}}$ from a conditional Wishart posterior distribution,
4. Sample $\beta \mid D_T, \pi, \phi_{-\beta}$ from a conditional Normal posterior distribution,
5. Sample $\Sigma_{\varepsilon\varepsilon}^{-1} \mid D_T, \pi, \phi_{-\Sigma_{\varepsilon\varepsilon}}$ from a conditional Wishart posterior distribution,
6. Sample $\pi \mid \phi$ using Carter and Kohn (1994),

7. Go to step 2 and continue until you reach a total of 120,000 iterations.

We now describe the prior distributions, the posterior distributions and the algorithm of Carter and Kohn (1994).

B.4 Priors

We now describe the priors on the parameters of the predictive system in Equations (A.9)–(A.11). The prior parameters for the benchmark priors are described in Section 4.1, whereas the prior parameters on the alternative priors in Section 5.1.

- β is a vector consisting of a_i, b_i, θ_i, and γ_i for $i = \{1, 2, 3\}$, and follows a multivariate normal distribution as

  $$\beta \sim \mathcal{N}\left(\underline{b}, \underline{B}\right),$$

  where $\underline{b}$ is the vector of prior means and $\underline{B}$ is the prior covariance matrix with off-diagonal elements set to zero. The prior density for β is specified as

  $$p(\beta) \propto \exp\left[-\frac{1}{2}(\beta - \underline{b})'\underline{B}^{-1}(\beta - \underline{b})\right] \cdot \mathbb{1}_{\beta \in \mathbb{R}},$$

  where $\mathbb{1}_{\beta \in \mathbb{R}}$ represents the acceptance region. In this case, the acceptance region is unbounded for a_i, b_i, and θ_i, while γ_i is constrained as $|\gamma_i| < 1$ to ensure stationarity.

- φ is a vector comprising δ_i, the elements of $\Sigma_{\eta u}$ (the covariances between unexpected returns and shocks to unobservable predictors), and the elements of $\Sigma_{\eta v}$ (the covariances between shocks to unobservable predictors and shocks to observable predictors). We specify a multivariate normal distribution as

  $$\varphi \sim \mathcal{N}\left(\underline{g}, \underline{G}\right)$$

  where $\underline{g}$ is the vector of prior means and $\underline{G}$ is the diagonal prior covariance matrix with off-diagonal elements set to zero. The elements of $\Sigma_{\eta u}$ have prior means and variances so that the implied prior on $\rho_{\eta_i u_i}$ has a given mean and a probability mass within a given interval. The elements of $\Sigma_{\eta v}$ have always prior means equal to zero and prior variances equal to one. The prior density for φ is specified as

$$p(\varphi) \propto \exp\left[-\frac{1}{2}(\varphi - \underline{g})'\underline{G}^{-1}(\varphi - \underline{g})\right] \cdot \mathbb{1}_{\varphi \in \mathbb{R}},$$

where $\mathbb{1}_{\varphi \in \mathbb{R}}$ represents the acceptance region. In this case, the acceptance region for δ_i is constrained as $0 < \delta_i < 1$ to ensure stationarity and positive correlation with the predicted return. Additionally, the acceptance region for the implied correlations $\rho_{\eta_i v_i}$ and $\rho_{\eta_i u_i}$ is between $[-1, 1]$.

- $\Sigma_{\varepsilon\varepsilon}^{-1}$ follows a Wishart distribution as

$$\Sigma_{\varepsilon\varepsilon}^{-1} \sim \mathcal{W}\left(\underline{S}_{\varepsilon}^{-1}, \underline{s}_{\varepsilon}\right)$$

where $\underline{S}_{\varepsilon}^{-1}$ is a positive definite prior scale matrix with off-diagonal element set to zero, and $\underline{s}_{\varepsilon}$ is the prior degree of freedom. This is equivalent to saying that $\Sigma_{\varepsilon\varepsilon} \sim \mathcal{W}^{-1}\left(\underline{S}_{\varepsilon}, \underline{s}_{\varepsilon}\right)$ and $\mathbb{E}\left(\Sigma_{\varepsilon\varepsilon}\right) = \underline{S}_{\varepsilon}^{-1}/(\underline{s}_{\varepsilon} - 2m - 1)$ with $\underline{s}_{\varepsilon} > 2m + 1$. We set $\underline{S}_{\varepsilon}^{-1}$ equal to the sample least-square estimate of $\Sigma_{\varepsilon\varepsilon}$, whereas $\underline{s}_{\varepsilon} = 2m + 2$. The prior density for $\Sigma_{\varepsilon\varepsilon}^{-1}$ is specified as

$$p(\Sigma_{\varepsilon\varepsilon}^{-1}) \propto \left|\Sigma_{\varepsilon\varepsilon}^{-1}\right|^{\frac{(\underline{s}_{\varepsilon} - 2m - 1)}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\Sigma_{\varepsilon\varepsilon}^{-1}\underline{S}_{\varepsilon}\right)\right].$$

- $\Sigma_{\zeta\zeta}^{-1}$ follows a Wishart distribution as

$$\Sigma_{\zeta\zeta}^{-1} \sim \mathcal{W}\left(\underline{S}_{\zeta}^{-1}, \underline{s}_{\zeta}\right),$$

where $\underline{S}_{\zeta}^{-1}$ is a positive definite prior scale matrix with off-diagonal element set to zero, and $\underline{s}_{\zeta}$ is the prior degree of freedom. Since $\Sigma_{\zeta\zeta} \sim \mathcal{W}^{-1}\left(\underline{S}_{\zeta}, \underline{s}_{\zeta}\right)$, we have that $E\left(\Sigma_{\zeta\zeta}\right) = \underline{S}_{\zeta}/(\underline{s}_{\zeta} - m - 1)$ with $\underline{s}_{\zeta} > m + 1$. We set $\underline{S}_{\zeta}$ equal to an identity matrix and

$\underline{s}_\zeta = m + 2$. The prior density for $\Sigma_{\zeta\zeta}^{-1}$ is specified as

$$p(\Sigma_{\zeta\zeta}^{-1}) \propto \left|\Sigma_{\zeta\zeta}^{-1}\right|^{\frac{(\underline{s}_\zeta - m - 1)}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\Sigma_{\zeta\zeta}^{-1}\underline{S}_\zeta\right)\right].$$

- π_0 follows a normal distribution as

$$\pi_0 \sim \mathcal{N}\left(b_0, Q_0\right),$$

where b_0 is the vector of prior means and Q_0 is a diagonal prior covariance matrix with off-diagonal elements set to zero. We set the prior means equal to zero and the prior covariance matrix equal to an identity matrix. The density of π_0 is

$$p(\pi_0) \propto \exp\left[-\frac{1}{2}(\pi_0 - b_0)' Q_0^{-1}(\pi_0 - b_0)\right].$$

B.5 Posterior Distributions

The joint posterior distribution is defined as

$$p\left(\phi, \pi \mid D_T\right) \propto p\left(D_T \mid \phi, \pi\right) \times p(\pi \mid \phi) \times p(\phi) \tag{A.16}$$

where the likelihood $p\left(D_T \mid \phi, \pi\right)$ is defined as

$$\begin{aligned} p\left(D_T \mid \phi, \pi\right) &= \prod\nolimits_{t=1}^{T} p(y_t \mid \pi_t, \phi) \\ &\propto \left|\Sigma_{\varepsilon\varepsilon}\right|^{-\frac{T}{2}} \exp\left[-\frac{1}{2}\sum\nolimits_{t=1}^{T}\left(y_t - X_{t-1}\beta - Z\pi_{t-1}\right)' \Sigma_{\varepsilon\varepsilon}^{-1}\left(y_t - X_{t-1}\beta - Z\pi_{t-1}\right)\right], \end{aligned}$$

the joint prior distribution $p(\phi)$ of the unknown parameters is simply

$$p\left(\phi\right) = p\left(\beta\right) p\left(\varphi\right) p\left(\Sigma_{\varepsilon\varepsilon}^{-1}\right) p\left(\Sigma_{\zeta\zeta}^{-1}\right),$$

and the prior distribution $p(\pi \mid \phi)$ of the state vector π is

$$\begin{aligned} p(\pi \mid \phi) &= \prod_{t=1}^{T} p(\pi_t \mid \pi_{t-1}, \phi) \\ &\propto |\Sigma_{\zeta\zeta}|^{-\frac{T}{2}} \exp\left[-\frac{1}{2}\sum_{t=1}^{T}(\pi_t - N_{t-1}\varphi)' \Sigma_{\zeta\zeta}^{-1}(\pi_t - N_{t-1}\varphi)\right]. \end{aligned}$$

While the joint posterior distribution does not take a convenient form, the conditional distributions are easy to derive. We now show the conditional posterior distributions.

B.5.1 Conditional Posterior of β

Start from the joint posterior distribution in Equation (A.16), set $\widetilde{y}_t = y_t - Z\pi_{t-1}$ and write the conditional posterior distribution of β as

$$\begin{aligned} p(\beta \mid D_T, \pi, \phi_{-\beta}) &\propto p(D_T \mid \phi, \pi) \times p(\beta) \\ &\propto \exp\left\{-\frac{1}{2}\begin{bmatrix} \sum_t (\widetilde{y}_t - X_{t-1}\beta)' \Sigma_{\varepsilon\varepsilon}^{-1}(\widetilde{y}_t - X_{t-1}\beta) \\ +(\beta - \underline{b})'\underline{B}^{-1}(\beta - \underline{b}) \end{bmatrix}\right\} \cdot \mathbb{1}_{\beta\in\mathbb{R}} \\ &\propto \exp\left\{-\frac{1}{2}\begin{bmatrix} \sum_t \widetilde{y}_t'\Sigma_{\varepsilon\varepsilon}^{-1}\widetilde{y}_t - 2\sum_t \beta' X_{t-1}'\Sigma_{\varepsilon\varepsilon}^{-1}\widetilde{y}_t \\ +\sum_t \beta' X_{t-1}'\Sigma_{\varepsilon\varepsilon}^{-1}X_{t-1}\beta \\ +\beta'\underline{B}^{-1}\beta - 2\underline{b}'\underline{B}^{-1}\beta + \underline{b}'\underline{B}^{-1}\underline{b} \end{bmatrix}\right\} \cdot \mathbb{1}_{\beta\in\mathbb{R}}. \end{aligned}$$

Remove the terms that do not contain β, set $\overline{B} = (\underline{B}^{-1} + \sum_t X_{t-1}'\Sigma_{\varepsilon\varepsilon}^{-1}X_{t-1})^{-1}$ and $\overline{b} = \overline{B}(\underline{B}^{-1}\underline{b} + \sum_t X_{t-1}'\Sigma_{\varepsilon\varepsilon}^{-1}\widetilde{y}_t)$, and obtain

$$p(\beta \mid D_T, \pi, \phi_{-\beta}) \propto \exp\left\{-\tfrac{1}{2}\left[(\beta - \overline{b})'\overline{B}^{-1}(\beta - \overline{b})\right]\right\} \cdot \mathbb{1}_{\beta\in\mathbb{R}}.$$

It follows that β has a conditional (truncated) normal posterior distribution

$$\beta \mid D_T, \pi, \phi_{-\beta} \sim \mathcal{N}\left(\overline{b}, \overline{B}\right) \cdot \mathbb{1}_{\beta\in\mathbb{R}},$$

with

$$\overline{B} = \left(\underline{B}^{-1} + \sum\nolimits_t X'_{t-1} \Sigma_{\varepsilon\varepsilon}^{-1} X_{t-1}\right)^{-1}$$

and

$$\bar{b} = \overline{B}\left[\underline{B}^{-1}\underline{b} + \sum\nolimits_t X'_{t-1} \Sigma_{\varepsilon\varepsilon}^{-1}\left(y_t - Z\pi_{t-1}\right)\right].$$

B.5.2 Conditional Posterior of $\Sigma_{\varepsilon\varepsilon}^{-1}$

Start from the joint posterior distribution, set $\widetilde{y}_t = y_t - X_{t-1}\beta - Z\pi_{t-1}$ and write the conditional posterior distribution of $\Sigma_{\varepsilon\varepsilon}^{-1}$ as

$$\begin{aligned} p\left(\Sigma_{\varepsilon\varepsilon}^{-1} \mid D_T, \pi, \phi_{-\Sigma_{\varepsilon\varepsilon}}\right) &\propto \left|\Sigma_{\varepsilon\varepsilon}\right|^{-\frac{T}{2}} \exp\left[-\frac{1}{2}\sum\nolimits_t \widetilde{y}'_t \Sigma_{\varepsilon\varepsilon}^{-1} \widetilde{y}_t\right] \times \left|\Sigma_{\varepsilon\varepsilon}^{-1}\right|^{\frac{(\underline{s}_\varepsilon - 2m-1)}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\Sigma_{\varepsilon\varepsilon}^{-1}\underline{S}_\varepsilon\right)\right] \\ &\propto \left|\Sigma_{\varepsilon\varepsilon}^{-1}\right|^{\frac{(T+\underline{s}_\varepsilon - 2m-1)}{2}} \exp\left[-\frac{1}{2}\sum\nolimits_t \widetilde{y}'_t \Sigma_{\varepsilon\varepsilon}^{-1} \widetilde{y}_t\right] \times \exp\left[-\frac{1}{2}\mathrm{tr}\left(\Sigma_{\varepsilon\varepsilon}^{-1}\underline{S}_\varepsilon\right)\right]. \end{aligned}$$

Using the properties of the trace operator,[5] we can rewrite the conditional posterior of $\Sigma_{\varepsilon\varepsilon}^{-1}$ as

$$\begin{aligned} p\left(\Sigma_{\varepsilon\varepsilon}^{-1} \mid D_T, \pi, \phi_{-\Sigma_{\varepsilon\varepsilon}}\right) &\propto \left|\Sigma_{\varepsilon\varepsilon}^{-1}\right|^{\frac{(T+\underline{s}_\varepsilon - 2m-1)}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Sigma_{\varepsilon\varepsilon}^{-1}\left(\sum\nolimits_t \widetilde{y}_t\widetilde{y}'_t + \underline{S}_\varepsilon\right)\right]\right\} \\ &\propto \left|\Sigma_{\varepsilon\varepsilon}^{-1}\right|^{\frac{(\overline{s}_\varepsilon - 2m-1)}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\overline{S}_\varepsilon\Sigma_{\varepsilon\varepsilon}^{-1}\right]\right\}, \end{aligned}$$

where $\overline{S}_\varepsilon = \underline{S}_\varepsilon + \sum_t \widetilde{y}_t\widetilde{y}'_t$ and $\overline{s}_\varepsilon = T + \underline{s}_\varepsilon$. It follows that $\Sigma_{\varepsilon\varepsilon}^{-1}$ has a conditional posterior Wishart distribution as

$$\Sigma_{\varepsilon\varepsilon}^{-1} \mid D_T, \pi, \phi_{-\Sigma_{\varepsilon\varepsilon}} \sim \mathcal{W}\left(\overline{S}_\varepsilon^{-1}, \overline{s}_\varepsilon\right),$$

[5] Recall that i) the trace of a scalar is the scalar itself, i.e., $\mathrm{tr}(a) = a$; ii) the trace operator is invariant under cyclic permutation, i.e., $\mathrm{tr}(AB) = \mathrm{tr}(BA)$; and iii) the trace is linear mapping, i.e., $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$.

with

$$
\begin{aligned}
\overline{S}_{\varepsilon} &= \underline{S}_{\varepsilon} + \sum\nolimits_{t} \left(y_t - X_{t-1}\beta - Z\pi_{t-1}\right)\left(y_t - X_{t-1}\beta - Z\pi_{t-1}\right)' \\
\overline{s}_{\varepsilon} &= T + \underline{s}_{\varepsilon}.
\end{aligned}
$$

B.5.3 Conditional Posterior of φ

Start from the joint posterior distribution and write the conditional posterior of φ as

$$
\begin{aligned}
p\left(\varphi \mid D_T, \pi, \phi_{-\varphi}\right) \quad &\propto \quad \exp\left\{-\frac{1}{2}\left[\begin{array}{c} \sum_t \left(\pi_t - N_{t-1}\varphi\right)' \Sigma_{\zeta\zeta}^{-1}\left(\pi_t - N_{t-1}\varphi\right) \\ +(\varphi-\underline{g})'\underline{G}^{-1}(\varphi-\underline{g}) \end{array}\right]\right\} \cdot \mathbb{1}_{\varphi\in\mathbb{R}} \\
&\propto \quad \exp\left\{-\frac{1}{2}\left[\begin{array}{c} \sum_t \pi_t'\Sigma_{\zeta\zeta}^{-1}\pi_t - 2\sum_t \pi_t'\Sigma_{\zeta\zeta}^{-1}N_{t-1}\varphi \\ +\sum_t \varphi' N_{t-1}'\Sigma_{\zeta\zeta}^{-1}N_{t-1}\varphi \\ +\varphi'\underline{G}^{-1}\varphi - 2\underline{g}'\underline{G}^{-1}\varphi + \underline{g}'\underline{G}^{-1}\underline{g} \end{array}\right]\right\} \cdot \mathbb{1}_{\varphi\in\mathbb{R}}.
\end{aligned}
$$

Remove the terms that do not contain φ, set $\overline{G} = (\underline{G}^{-1} + \sum_t N_{t-1}'\Sigma_{\zeta\zeta}^{-1}N_{t-1})^{-1}$ and $\overline{g} = \overline{G}\left(\underline{G}^{-1}\underline{g} + \sum_t N_{t-1}'\Sigma_{\zeta\zeta}^{-1}\pi_t\right)$, and obtain

$$
p\left(\varphi \mid D_T, \pi, \phi_{-\varphi}\right) \quad \propto \exp\left\{-\tfrac{1}{2}\left[(\varphi-\overline{g})'\overline{G}^{-1}(\varphi-\overline{g})\right]\right\} \cdot \mathbb{1}_{\varphi\in\mathbb{R}}.
$$

It follows that φ has a conditional posterior normal distribution as

$$
\varphi \mid D_T, \pi, \phi_{-\varphi} \sim \mathcal{N}\left(\overline{g}, \overline{G}\right) \cdot \mathbb{1}_{\varphi\in\mathbb{R}},
$$

where

$$
\begin{aligned}
\overline{G} &= \left(\underline{G}^{-1} + \sum\nolimits_t N_{t-1}'\Sigma_{\zeta\zeta}^{-1}N_{t-1}\right)^{-1} \\
\overline{g} &= \overline{G}\left(\underline{G}^{-1}\underline{g} + \sum\nolimits_t N_{t-1}'\Sigma_{\zeta\zeta}^{-1}\pi_t\right).
\end{aligned}
$$

B.5.4 Conditional Posterior of $\Sigma_{\zeta\zeta}^{-1}$

Start from the joint posterior distribution, set $\widetilde{\pi}_t = \pi_t - N_{t-1}\varphi$ and write the conditional posterior of $\Sigma_{\zeta\zeta}^{-1}$ as

$$\begin{aligned}
p\left(\Sigma_{\zeta\zeta}^{-1} \mid D_T, \pi, \phi_{-\Sigma_{\zeta\zeta}}\right) \;&\propto\; |\Sigma_{\zeta\zeta}|^{-\frac{T}{2}} \exp\left[-\frac{1}{2}\sum\nolimits_t \widetilde{\pi}_t' \Sigma_{\zeta\zeta}^{-1}\widetilde{\pi}_t\right] \times \left|\Sigma_{\zeta\zeta}^{-1}\right|^{\frac{(\underline{s}_\zeta - m-1)}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\Sigma_{\zeta\zeta}^{-1}\underline{S}_\zeta\right)\right] \\
&\propto\; \left|\Sigma_{\zeta\zeta}^{-1}\right|^{\frac{(T+\underline{s}_\zeta - m-1)}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Sigma_{\zeta\zeta}^{-1}\left(\sum\nolimits_t \widetilde{\pi}_t\widetilde{\pi}_t' + \underline{S}_\zeta\right)\right]\right\} \\
&\propto\; \left|\Sigma_{\zeta\zeta}^{-1}\right|^{-\frac{(\overline{s}_\zeta - m-1)}{2}} \exp\left[-\frac{1}{2}\mathrm{tr}\left(\Sigma_{\zeta\zeta}^{-1}\overline{S}_\zeta\right)\right].
\end{aligned}$$

It follows that $\Sigma_{\zeta\zeta}^{-1}$ has a conditional posterior Wishart distribution as

$$\Sigma_{\zeta\zeta}^{-1} \mid D_T, \pi, \phi_{-\Sigma_{\zeta\zeta}} \sim \mathcal{W}\left(\overline{S}_\zeta^{-1}, \overline{s}_\zeta\right),$$

where

$$\begin{aligned}
\overline{S}_\zeta &= \underline{S}_\zeta + \sum\nolimits_t (\pi_t - N_{t-1}\varphi)(\pi_t - N_{t-1}\varphi)' \\
\overline{s}_\zeta &= T + \underline{s}_\zeta.
\end{aligned}$$

B.6 Sampling $\pi \mid \phi$ using Carter and Kohn (1994)

We sample the vector π in one block from the full conditional posterior distribution $p(\pi \mid D_T, \phi)$ using the *forward filtering, backward sampling* method of Carter and Kohn (1994). We will omit ϕ throughout this section for simplicity.

B.6.1 Forward Filtering

Recall the state space system in Equations (A.9)–(A.11),

$$
\begin{aligned}
y_t &= X_{t-1}\beta + Z\pi_{t-1} + \varepsilon_t \\
\pi_t &= \delta\pi_{t-1} + \eta_t,
\end{aligned}
$$

with

$$
\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \overset{i.i.d.}{\sim} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{\varepsilon\varepsilon} & \Sigma'_{\eta\varepsilon} \\ \Sigma_{\eta\varepsilon} & \Sigma_{\eta\eta} \end{bmatrix} \right).
$$

Let $D_t = \{y_t, D_{t-1}\}$ be the information set at time t. Forward filtering consists of the following steps:

a) Initial condition at time $t-1$

$$
\pi_{t-1} \mid D_{t-1} \sim \mathcal{N}\left(b_{t-1}, Q_{t-1}\right).
$$

b) Prior at time t

$$
\pi_t \mid D_{t-1} \sim \mathcal{N}\left(a_t, P_t\right),
$$

where

$$
\begin{aligned}
a_t &= E\left(\pi_t \mid D_{t-1}\right) = \delta E\left(\pi_{t-1} \mid D_{t-1}\right) = \delta b_{t-1} \\
P_t &= Var\left(\pi_t \mid D_{t-1}\right) = \delta Var\left(\pi_{t-1} \mid D_{t-1}\right)\delta' + \Sigma_{\eta\eta} = \delta Q_{t-1}\delta' + \Sigma_{\eta\eta}.
\end{aligned}
$$

c) Prediction at time t

$$
y_t \mid D_{t-1} \sim \mathcal{N}\left(f_t, S_t\right),
$$

where

$$f_t = E\left(y_t \mid D_{t-1}\right) = X_{t-1}\beta + ZE\left(\pi_{t-1} \mid D_{t-1}\right) = X_{t-1}\beta + Zb_{t-1}$$
$$S_t = Var\left(y_t \mid D_{t-1}\right) = ZVar\left(\pi_{t-1} \mid D_{t-1}\right)Z' + \Sigma_{\varepsilon\varepsilon} = ZQ_{t-1}Z' + \Sigma_{\varepsilon\varepsilon}.$$

d) Joint distribution at time t

$$\begin{bmatrix} y_t \\ & \mid D_{t-1} \\ \pi_t \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} f_t \\ a_t \end{bmatrix}, \begin{bmatrix} S_t & G_t \\ G_t' & P_t \end{bmatrix}\right),$$

where

$$\begin{aligned} G_t &= Cov\left(y_t, \pi_t \mid D_{t-1}\right) \\ &= ZVar\left(\pi_{t-1} \mid D_{t-1}\right)\delta' + Cov\left(\varepsilon_t, \eta_t \mid D_{t-1}\right) \\ &= ZQ_{t-1}\delta' + \Sigma_{\eta\varepsilon}'. \end{aligned}$$

e) Posterior at time t

$$\pi_t \mid D_t \sim \mathcal{N}\left(b_t, Q_t\right),$$

where

$$b_t = E\left(\pi_t \mid y_t, D_{t-1}\right) = a_t + G_t'S_t^{-1}\left(y_t - f_t\right)$$
$$Q_t = Var\left(\pi_t \mid y_t, D_{t-1}\right) = P_t - G_t'S_t^{-1}G_t.$$

The posterior hyperparameters b_t and Q_t are easily derived since $\pi_t \mid D_t$ is equivalent to $\pi_t \mid y_t, D_{t-1}$. Using the joint distribution of y_t and π_t presented above, it is easy to obtain the conditional distributions from a multivariate normal distribution.

B.6.2 Backward Filtering

The backward sampling method builds on the following Markov property

$$p\left(\xi_1,\ldots,\xi_T \mid D_T\right) = p\left(\xi_T \mid D_T\right) p\left(\xi_{T-1} \mid \xi_T, D_{T-1}\right) \times \cdots \times p\left(\xi_1 \mid \xi_2, D_1\right).$$

We sample π_T from $p\left(\pi_T \mid D_T\right)$ and then π_t from the conditional density $p\left(\xi_t \mid \xi_{t+1}, D_t\right)$ for $t = T-1,\ldots,1$, where $\xi_t = [y_t, \pi_t]'$. We derive the conditional density $p\left(\xi_t \mid \xi_{t+1}, D_t\right)$ as follows

$$\underbrace{\begin{bmatrix} y_t \\ \pi_t \end{bmatrix}}_{\xi_t} = \underbrace{\begin{bmatrix} 0 & Z \\ 0 & \delta \end{bmatrix}}_{M} \underbrace{\begin{bmatrix} y_{t-1} \\ \pi_{t-1} \end{bmatrix}}_{\xi_{t-1}} + \underbrace{\begin{bmatrix} 0 & \beta \\ 0 & 0 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 1 \\ X_{t-1} \end{bmatrix}}_{\Lambda_{t-1}} + \underbrace{\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix}}_{e_t}$$

and recall that

$$\xi_{t+1} \mid D_t \sim \mathcal{N}\left(\underbrace{\begin{bmatrix} f_{t+1} \\ a_{t+1} \end{bmatrix}}_{\overline{a}_{t+1}}, \underbrace{\begin{bmatrix} S_{t+1} & G_{t+1} \\ G'_{t+1} & P_{t+1} \end{bmatrix}}_{\overline{A}_{t+1}} \right)$$

and

$$\xi_t \mid D_t \sim \mathcal{N}\left(\underbrace{\begin{bmatrix} y_t \\ b_t \end{bmatrix}}_{\overline{b}_t}, \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & Q_t \end{bmatrix}}_{\overline{B}_t} \right).$$

The **joint density** is then straightforward to derive as

$$\begin{bmatrix} \xi_t & \\ & \mid D_t \\ \xi_{t+1} & \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \overline{b}_t \\ \overline{a}_{t+1} \end{bmatrix}, \begin{bmatrix} \overline{B}_t & \overline{G}_t \\ \overline{G}'_t & \overline{A}_{t+1} \end{bmatrix} \right),$$

where

$$\begin{aligned}
\overline{G}_t &= Cov\left(\xi_t, \xi_{t+1} \mid D_t\right) \\
&= Cov\left(\xi_t, M\xi_t + L\Lambda_t + e_{t+1} \mid D_t\right) \\
&= Var\left(\xi_t \mid D_t\right) M' \\
&= \overline{B}_t M'.
\end{aligned}$$

Hence, we have obtained

$$\xi_t \mid \xi_{t+1}, D_t \sim \mathcal{N}\left(h_t, H_t\right),$$

where

$$\begin{aligned}
h_t &= E\left(\xi_t \mid \xi_{t+1}, D_t\right) = \overline{b}_t + \overline{G}_t \overline{A}_{t+1}^{-1} \left(\xi_{t+1} - \overline{a}_{t+1}\right) \\
H_t &= Var\left(\xi_t \mid \xi_{t+1}, D_t\right) = \overline{B}_t - \overline{G}_t \overline{A}_{t+1}^{-1} \overline{G}_t'.
\end{aligned}$$

B.6.3 Summary

We now summarize the forward filtering and backward filtering algorithm as

a) Prediction Equations

$$\begin{aligned}
E\left(\pi_t \mid D_{t-1}\right) &: a_t = \delta b_{t-1} \\
Var\left(\pi_t \mid D_{t-1}\right) &: P_t = \delta Q_{t-1}' \delta + \Sigma_{\eta\eta} \\
E\left(y_t \mid D_{t-1}\right) &: f_t = X_{t-1}\beta + Z b_{t-1} \\
Var\left(y_t \mid D_{t-1}\right) &: S_t = Z Q_{t-1} Z' + \Sigma_{\varepsilon\varepsilon} \\
Cov\left(y_t, \pi_t \mid D_{t-1}\right) &: G_t = Z Q_{t-1} \delta + \Sigma_{\eta\varepsilon}'.
\end{aligned}$$

b) Updating Equations

$$E(\pi_t \mid y_t, D_{t-1}) : b_t = a_t + G_t' S_t^{-1}(y_t - f_t)$$
$$Var(\pi_t \mid y_t, D_{t-1}) : Q_t = P_t - G_t' S_t^{-1} G_t.$$

c) Sample π_T^* from $p(\pi_T \mid D_T, \phi)$

$$\pi_T^* \sim \mathcal{N}(b_T, Q_T).$$

d) Sample π_t^* from $p(\pi_t \mid \pi_{t+1}, D_t, \phi)$ starting from $t = T-1, \ldots, 1$

$$\pi_t^* \sim \mathcal{N}(h_{t,\pi}, H_{t,\pi}),$$

where

$$E(\xi_t \mid \xi_{t+1}, D_t) : h_t = \overline{b}_t + \overline{G}_t \overline{A}_{t+1}^{-1}(\xi_{t+1} - \overline{a}_{t+1})$$
$$Var(\xi_t \mid \xi_{t+1}, D_t) : H_t = \overline{B}_t - \overline{G}_t \overline{A}_{t+1}^{-1} \overline{G}_t',$$

with

$$\overline{b}_t = \begin{bmatrix} y_t \\ b_t \end{bmatrix}, \quad \overline{G}_t = \overline{B}_t M', \quad \overline{B}_t = \begin{bmatrix} 0 & 0 \\ 0 & Q_t \end{bmatrix}, \quad M = \begin{bmatrix} 0 & Z \\ 0 & \delta \end{bmatrix},$$

$$\overline{A}_{t+1} = \begin{bmatrix} S_{t+1} & G_{t+1} \\ G_{t+1}' & P_{t+1} \end{bmatrix}, \quad \overline{a}_{t+1} = \begin{bmatrix} f_{t+1} \\ a_{t+1} \end{bmatrix}, \quad \xi_{t+1} = \begin{bmatrix} y_{t+1} \\ \pi_{t+1} \end{bmatrix}.$$

The generated vector $[\pi_1^*, \ldots, \pi_T^*]'$ corresponds to a random draw from $p(\pi_1, \ldots, \pi_T \mid D_T)$.

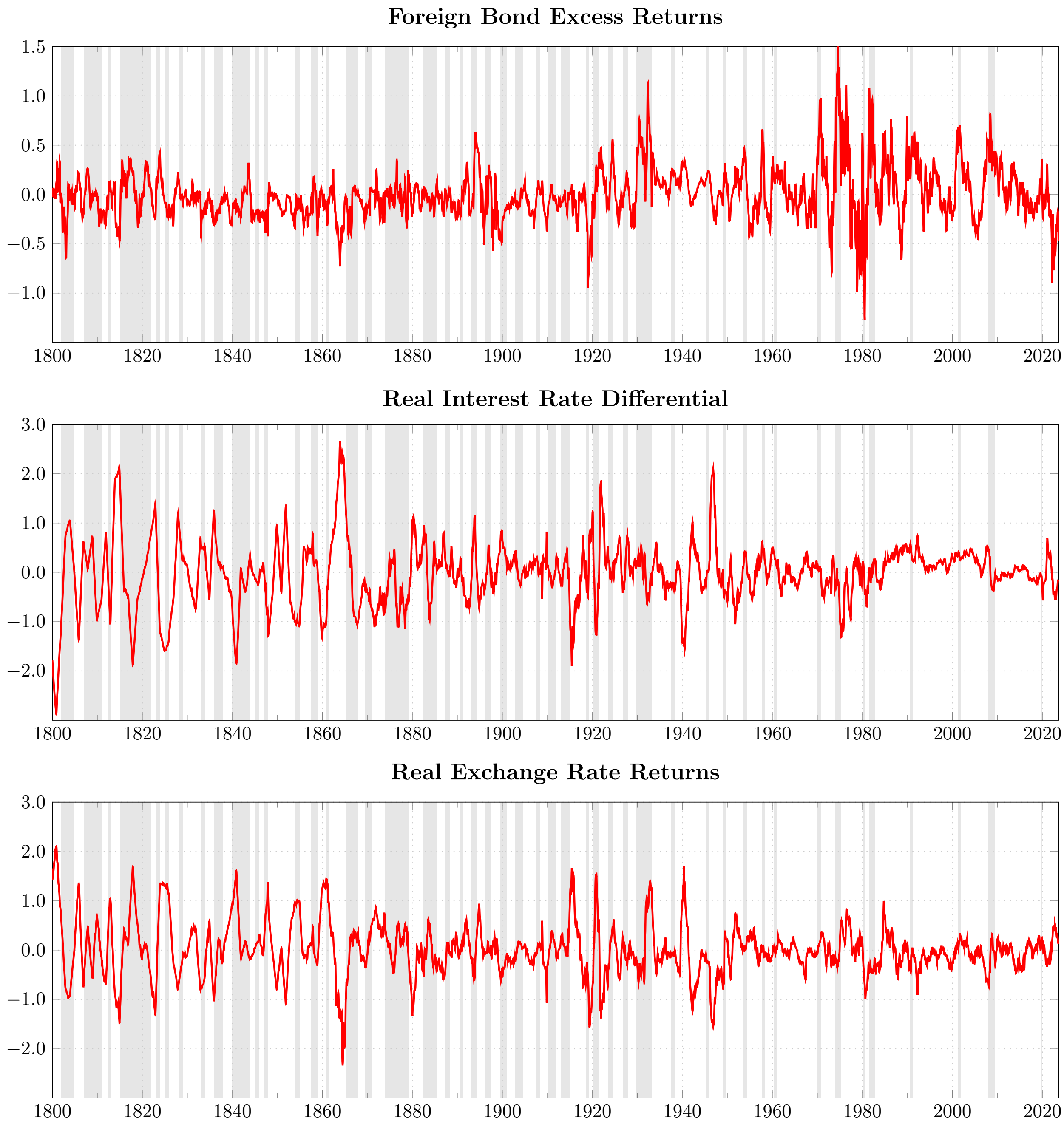

Figure A.1: Unobservable Predictors

This figure plots the unobservable predictors of the predictive system in Equations (2)–(5). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. Each unobservable predictor is sampled in one block using the *forward filtering, backward sampling* approach of Carter and Kohn (1994). The sample includes monthly data from December 1799 to December 2023 for the UK relative to the US, sourced from Global Financial Data. The shaded areas are NBER recession periods from the FRED database.

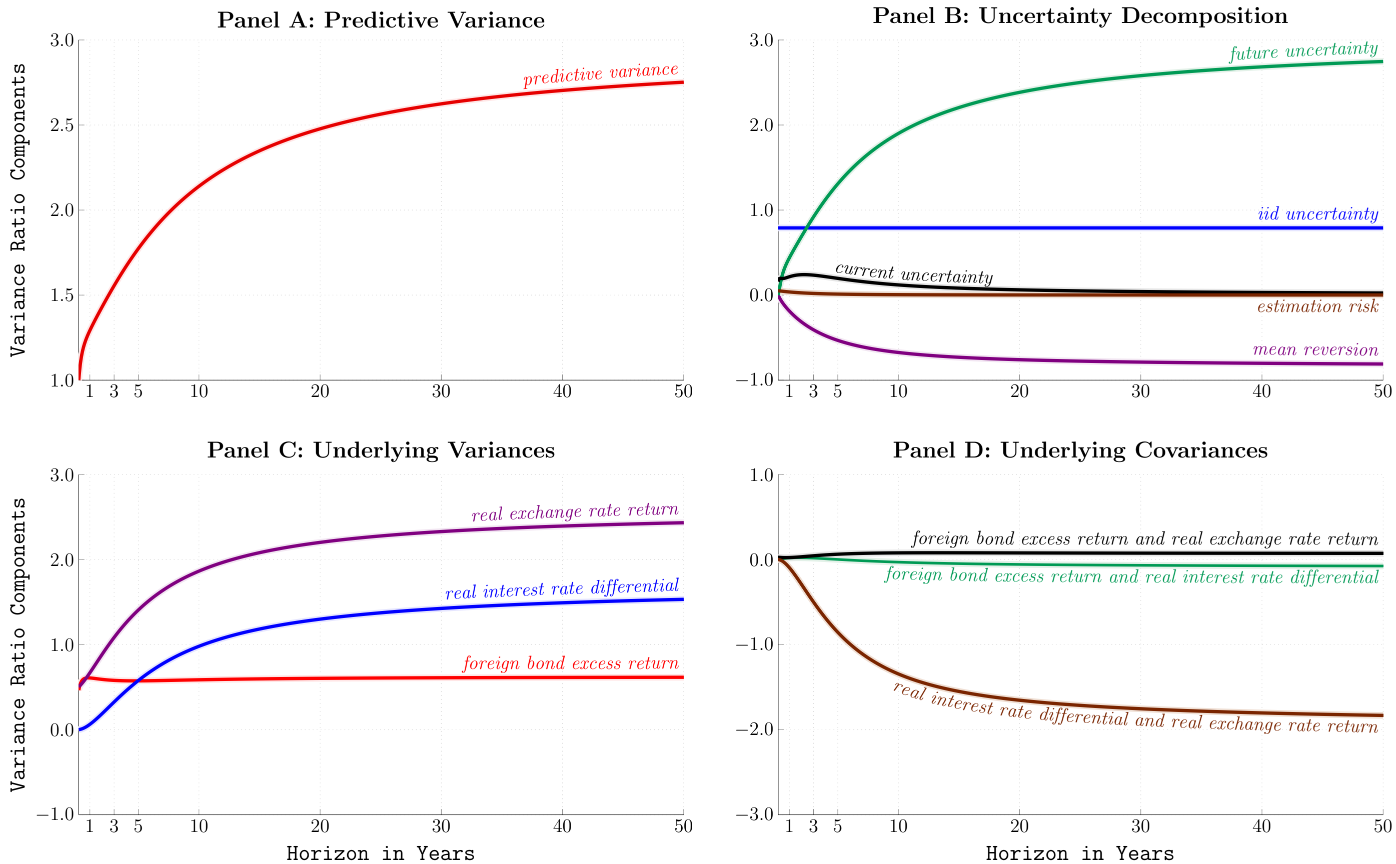

Figure A.2: Predictive Variance of an Equally-Weighted Portfolio

This figure presents the predictive variance of the excess return for a strategy buying 10-year constant maturity bonds in foreign currency with equal weights while borrowing at the 3-month interest rate in US dollars. Panel A presents the predictive variance of the strategy, whereas Panel B breaks down the predictive variance into *iid uncertainty*, *mean reversion*, and *future expected return uncertainty* via Proposition 1, and the latter into *current expected return uncertainty* and *estimation risk* via Proposition 2. Panels C and D plot the underlying variances and covariances (each counted twice) of foreign bond excess return, real interest rate differential, and real exchange rate return, , respectively. All components are scaled by the predictive variance of a one-period excess return to ease the comparison across horizons as in Equation (17). Parameter estimates are from a Gibbs sampling algorithm with 100,000 iterations (following a burn-in of 20,000) and retaining every tenth draw. The sample includes monthly data from December 1799 to December 2023 for Australia, Canada, Germany, Japan, New Zealand, Norway, Sweden, Switzerland, and the UK relative to the US, sourced from Global Financial Data.