Macro Strikes Back: Term Structure of Risk Premia*

Svetlana Bryzgalova† Jiantao Huang‡ Christian Julliard§

February 2025

Abstract

We develop a unified framework to study the term structure of risk premia of nontradable factors. Our method delivers level and time variation of risk premia, uncovers their propagation mechanism, and is robust to misspecification and weak identification. Most macroeconomic factors are weakly identified at the quarterly frequency, but have increasing (unconditional) term structures with large risk premia at business cycle horizons. Moreover, macro risk premia are strongly time-varying and countercyclical. Our framework also recovers the term structure of forward equity yields. We show that it is strongly countercyclical and closely matches the observed values implied by dividend strip data.

Keywords: Macro-finance, asset pricing, risk premia, linear factor models, Bayesian inference, term structures.

JEL Classification Codes: C11, C58, E27, E44, G12, G17.

*Any errors or omissions are the responsibility of the authors. For helpful comments, discussions and suggestions, we thank Nina Boyarchenko, Mikhail Chernov, Richard Crump, Magnus Dahlquist, Marco Del Negro, Ian Dew-Becker, Xiang Fang, Marcelo Fernandes, Domenico Giannone, Bernardo Guimaraes, Gabriele La Spada, Ian Martin, Anna Mikusheva, Lasse Pedersen, Chenjie Xu, and seminar and conference participants at Copenhagen Business School, CUHK, FGV São Paulo, FinEML, London School of Economics, Renmin University of China, Duke/UNC 2024 Asset Pricing conference, SoFiE 2024 conference, CICF 2024, NBER summer institute 2024, and EEA-ESEM 2024.

†Department of Finance, London Business School; sbryzgalova@london.edu

‡Faculty of Business and Economics, The University of Hong Kong, huangjt@hku.hk

§Department of Finance, FMG, and SRC, London School of Economics, and CEPR; c.julliard@lse.ac.uk.

1 Introduction

Macroeconomic risk, often related to technology, consumption, or intermediary capital, is at the heart of most equilibrium-based asset pricing models. Yet reliable detection of macroeconomic risk premia remains elusive: 1) different time horizons often provide drastically different estimates of the priced risk, 2) most empirical models are widely known to be misspecified, calling for methods robust to the nature and number of risk factors, and 3) the weak contemporaneous link between macroeconomic factors and asset returns often leads to model parameters being weakly identified at best, causing a fundamental inference problem. All of these issues contribute to the empirical macrofinance disconnect.

We propose a new estimation framework that addresses all of the above. Unlike any existing approach, it produces not only reliable risk premia estimates but also their whole term structure in an internally consistent framework. Our method, which leverages the fact that many nontradable factors are persistent, relies on three key ingredients: 1) the moving average (MA) representation of the persistent component of the factor, driven by either priced or non-priced shocks, 2) an approximate factor structure for a wide cross-section of asset returns, which recovers priced shocks and is robust to model misspecification (Chamberlain and Rothschild (1983) and Giglio and Xiu (2021)), and 3) the hierarchical Bayesian inference method of Bryzgalova, Huang, and Julliard (2023), which recovers both time series and cross-sectional properties of risk factors, and is by design robust to weak identification.

Our framework accurately identifies not only the joint comovement between nontradable factors and asset returns but also their propagation mechanism, and hence recovers the whole *term structure of risk premia*. As we show, the latter is crucial in assessing the role of macroeconomic risks in asset returns. We find that many macroeconomic variables (e.g., industrial production, consumption, and GDP growth) have increasing unconditional term structures and carry large and significant risk premia at business cycle frequencies (two–three years). Furthermore, conditionally, their risk premia are strongly time-varying and countercyclical.

Our findings are not a simple byproduct of factor persistence. We find that similarly persistent risk factors can have increasing (liquidity of Pástor and Stambaugh (2003)), flat (intermediary factor of He et al. (2017)), or decreasing term structures (VIX), or no significant risk premia at all (capital share growth of Lettau et al. (2019)). As we show, risk premia over different horizons can be directly mapped into per-period average returns of factor-mimicking portfolios hedging multi-horizon priced innovations. The economic magnitude of our findings is striking: At business cycle frequencies, risk premia carried by, for example, industrial production, GDP, and consumption, are as large as that of the market. Crucially, all of our results are not based on ad-hoc frequency-based procedures, but are instead fully determined by the structural parameters of the model.

Our framework provides sharp identification of risk premia and reliable inference and is rooted in economic theory: Equilibrium asset prices are jump variables. Hence, news about current and *future* priced states should immediately be reflected into prices, albeit they might manifest in nontradable variables only with delay (see, e.g., Hansen et al. (2008)). Similar to Giglio and Xiu (2021), we leverage the fact that, while the actual drivers of asset returns are identifiable only up to a rotation, conditional and unconditional risk premia of observable factors are not affected by this issue and can, therefore, be reliably recovered from the data. As a result, our estimator is robust to the omitted variable bias, measurement error, and weak identification. Contrary to the existing literature, our method allows for the joint modeling of factor and return dynamics over different horizons, providing coherent insights into the whole term structure of risk premia. Tackling an inference problem in this setup would be challenging, if not infeasible, in frequentist estimation. Instead, we develop a simple Gibbs algorithm for Bayesian posterior sampling, with all the conditional posterior distributions available in closed form. Thus, we deliver not only point estimates of risk premia and deep model parameters but also valid credible intervals for all the objects of interest.

It is widely known that some nontradable factors have higher exposure to asset returns at longer horizons. (See, e.g., Jagannathan and Wang (2007), Cohen et al. (2009), and Hansen et al. (2008)) We uncover the mechanism generating this phenomenon and show that, in these cases, risk premia at different horizons are driven by the *same* priced innovations that slowly propagate through the nontradable risk factor. As a result, we also explain why many risk factors are statistically weak at quarterly frequency, yet become strongly identified at longer horizons. Consider GDP growth, for example. Although contemporaneous asset return shocks account for only 4% of the variation in GDP growth, they contribute to over 20% of its time series variation at business cycle frequencies. In such a case, the term structure of risk premia effectively boosts the signal-to-noise ratio of priced shocks in nontradable factors and sharply identifies the common priced component (which would normally be weak at best).

The empirical asset pricing literature has long recognized the persistent nature of many nontradable risk factors. As a result, researchers would usually first extract the AR(1) innovation from a factor and then proceed with measuring its risk premia via Fama-MacBeth (FM) regressions or the Generalized Method of Moments (GMM).[1] However, as we show, this common procedure fails to recover the true sources of priced risk. First, the conditional mean of the macroeconomic variable could follow a process different from AR(1). Second, the persistent component of the variable does not need to be driven only by priced shocks. As we show, AR(1) residuals do not recover actual priced innovations in many factors, leading to a significant bias in risk premia estimates. Our approach, rooted in the Wold decomposition, relies on the flexible MA representation of the risk factor. Like the local projection framework (Jordà, 2005), it recovers the impulse response functions of macro quantities to shocks spanned by financial markets *without* postulating a stringent (V)AR structure, hence avoiding the resulting fragility to misspecification (Olea et al., 2024). Different from these papers, our approach leverages the large cross-section of asset returns to efficiently separate priced and unpriced innovations and restore reliable inference on risk premia.

We use a large cross-section of 275 equity portfolios to estimate the term structures of risk premia of many nontradable risk factors. Contrary to the standard one-period inference, we find that a large part of the factors' conditional mean is driven by priced shocks, slowly propagating through the time series. Their overall dynamics display clear business cycle patterns and are common across many different macroeconomic factors.

Many risk factors are characterized by increasing term structures of risk premia. For example, while the risk premium of GDP growth is only 0.03 at the quarterly horizon, it increases to 0.20 at the three-year horizon and is strongly significant (while being spanned by the same shocks). Furthermore, the term structures of macro risk exhibit strong commonality in their business cycle behavior: The average level is strongly countercyclical, with moderate risk premia during expansion and significant increases during recessions.

We also observe factors commanding flat or downward-sloping unconditional term structures of risk premia. For example, the VIX risk premium is -0.13 at the monthly frequency, but its two-year counterpart is only -0.03. This observation is reassuring since the sign of the VIX risk premium, and its term structure, are mostly consistent with previous findings based on VIX derivatives (Eraker and Wu (2014), Dew-Becker et al. (2017), and Johnson (2017)). Intermediary factors (Adrian et al. (2014) and He et al. (2017)) carry significantly positive

[1]See He et al. (2017), Pástor and Stambaugh (2003), and Giglio and Xiu (2021), among others.

unconditional risk premia with an almost flat term structure. Furthermore, our estimate of risk premia for the intermediary factor of He et al. (2017) is rather close to the average return of its tradable version, further validating our findings.

We further investigate the connection between the term structure of risk premia and forward equity yields implied by the dividend strips data. Under joint lognormality, forward equity yields equal the difference between dividend risk premia and expected dividend growth. We show that our estimates of unconditional dividend risk premia between one- and five-year horizons are similar to those of Bansal et al. (2021). Note that we use a longer time series sample and an entirely different method from Bansal et al. (2021), so this consistency is affirmative.

To infer the term structure of forward equity yields, one needs an estimate of expected dividend growth. We rely on our MA representation to derive the conditional dynamics of dividend growth. However, we do not assume that the spanned MA component captures the entire dividend predictability (we allow for predictability of the unspanned component); hence, ex-ante, our formulation is not guaranteed to generate equity yields consistent with the observed data. Nevertheless, ex-post, our estimates of forward equity yields are strongly countercyclical and closely track the observed yields implied by the dividend strips data. Our model also generates a downward-sloping (upward-sloping) term structure of equity yields in economic downturns (expansions). The procyclical behaviours of the slopes of the dividend term structure are mainly driven by the time variation in expected dividend growth instead of time-varying dividend risk premia.

To further highlight the strength and robustness of our method, we conduct extensive simulations and study the empirical size and power of the procedure in detecting the persistent priced component of nontradable factors. We find that the Bayesian credible intervals provide proper posterior coverages of the pseudo-true risk premia. Moreover, we show that the MA-based approach is essential in identifying the priced component in persistent factors and leads to a significant boost in power. Finally, our Bayesian inference remains valid even in the presence of persistent yet weak risk factors.

The remainder of the paper is organized as follows. In the next subsection we review the most closely related literature and our contribution to it. Section 2 outlines our estimation framework and its properties, while Section 3 provides simulation evidence on the power of the method in realistically small samples. Section 4 presents our empirical findings, and Section 5 concludes. Additional results, proofs, derivations, and a detailed description of the data sources, are reported in the Internet Appendix.

Closely Related Literature

Our paper naturally relates to the inference on risk premia in linear factor models. As shown by the past literature (e.g., Kan and Zhang (1999a,b), Kleibergen (2009), and Kleibergen and Zhan (2015)), weak factors invalidate risk premia estimates and cross-sectional fit of traditional FM and GMM estimators. Several studies (e.g., Kan et al. (2013), Gospodinov et al. (2014, 2019), Bryzgalova (2015), Kleibergen and Zhan (2020), Anatolyev and Mikusheva (2022), and Bryzgalova et al. (2023)) propose methods that are robust to weak factors and misspecification. Giglio and Xiu (2021) further emphasize that standard estimators of risk premia are biased if some priced factors are omitted and propose a three-pass method to resolve the issue. Likewise, Giglio et al. (2023) propose a supervised principal component analysis (PCA) method to recover the risk premia of weak factors. Similarly, our method aligns with the literature that relates to PCA in asset pricing (e.g., Chamberlain and Rothschild (1983), Connor and Korajczyk (1986, 1988), Kozak et al. (2018, 2020), and Kelly et al. (2019)). Unlike them, we incorporate the dynamics of factors and returns to elicit the *entire* term structure of risk premia in an internally consistent manner. This additional dimension is economically meaningful since many variables, particularly macro variables, are significantly priced only at particular horizons.

Equilibrium macro-finance models have sharp and salient predictions for the term structures of risk premia of macro factors. For instance, as shown in Figure A1, the habit model of Campbell and Cochrane (1999) predicts flat term structures of risk premia for consumption and dividend growth.[2] Yet these same factors command upward-sloping term structures in the long-run risk model of Bansal and Yaron (2004). However, these predictions rely on ad hoc assumptions on cash flow dynamics and investors' preferences. To obtain model-free estimates, van Binsbergen et al. (2012) and van Binsbergen and Koijen (2017)) analyze traded dividend claims and observe a downward-sloping term structure of dividend risk, which contradicts the predictions of leading macro-finance models. Consequently, several equilibrium models (e.g., Belo et al. (2015), Hasler and Marfe (2016), Ai et al. (2018), and Kragt et al. (2020)) have been developed to explain this phenomenon. However, traded dividend strips data suffer from a short time series sample and liquidity concerns. Recent papers tackle these shortcomings by

[2]Calibration and derivation details can be found in Internet Appendix IA.4.

estimating either a regime-switching model (Bansal et al. (2021)) or an affine term structure model of expected returns and dividend growth (Giglio et al. (2023)). Both papers suggest an unconditionally upward-sloping and conditionally time-varying term structure of dividend risk – as we uncover for dividend growth risk premia. But crucially, our new method has much broader applicability than solely dividends, as it delivers the term structure of risk premia for *all* factors (traded and nontraded) of equilibrium models.

Our paper also connects to the large body of literature that emphasizes horizon-dependent risk premia (see, e.g., Chernov et al. (2021)). Extensive empirical evidence shows that consumption growth carries more significant premia at long horizons (Daniel and Marshall (1997), Parker and Julliard (2005), Jagannathan and Wang (2007), Hansen et al. (2008), Malloy et al. (2009), Ortu et al. (2013), Dew-Becker and Giglio (2016), and Bandi and Tamoni (2023)). In contrast, VIX (Eraker and Wu (2014), Dew-Becker et al. (2017), and Johnson (2017)) carry more sizable risk premia at short horizons. Our paper is motivated by these empirical facts and provides a much more extensive and robust investigation of the risk premia of more than 20 economic variables.

Finally, our paper is related to the recent developments of Bayesian econometrics in asset pricing (e.g., Barillas and Shanken (2018), Chib et al. (2020), Bryzgalova et al. (2023), and Avramov et al. (2023)). Unlike most papers, which emphasize Bayesian model selection and/or aggregation, we estimate the posterior credible intervals of the term structure of risk premia.

2 Theory and Method

This section describes our Bayesian framework for estimating factors' risk premia. We aim to test whether a (covariance-stationary) factor g_t, either tradable or nontradable, is priced in a large cross-section of test assets. Throughout our analysis, we consider log variables; that is, g_t is the log growth rate of G_t between time $t-1$ and t, where G_t can be, for example, the portfolio value, consumption, or production.

We denote the vector of log returns on N assets, in excess of the log risk-free rate (r_f), by $\boldsymbol{r_t} = (r_{1t}, \ldots, r_{Nt})^\top$. We further define the cumulative variable: $g_{t-1\to t+S} = \log(G_{t+S}) - \log(G_{t-1})$, which measures the multiperiod growth rate of G_t. Similarly, $\boldsymbol{r}_{t-1\to t+S}$ denote the cumulative log returns between time $t-1$ and $t+S$.

We assume a linear latent factor model for $\boldsymbol{r}_t$ driven by K systematic factors, as follows:

$$\boldsymbol{r}_t = \boldsymbol{\mu}_r + \boldsymbol{\beta}_{\tilde{v}}\tilde{\boldsymbol{v}}_t + \boldsymbol{w}_{rt}, \quad \tilde{\boldsymbol{v}}_t \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}_K, \boldsymbol{I}_K), \quad \boldsymbol{w}_{rt} \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}_N, \boldsymbol{\Sigma}_{wr}), \quad \tilde{\boldsymbol{v}}_t \perp \boldsymbol{w}_{rt}, \tag{1}$$

where $\tilde{\boldsymbol{v}}_t$ are K uncorrelated latent factors with loadings $\boldsymbol{\beta}_{\tilde{v}}$, $\boldsymbol{w}_{rt}$ are unpriced idiosyncratic errors, and $\boldsymbol{\mu}_r$ denote expected log excess returns. We relax the assumption of serially uncorrelated $\tilde{\boldsymbol{v}}_t$ in Section 2.2. We impose an approximate factor structure among asset returns, following Chamberlain and Rothschild (1983). Mathematically, the largest K eigenvalues of $\boldsymbol{r}_t$'s covariance matrix will explode as the number of assets goes to infinity (equivalently, the eigenvalues of $\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{\beta}_{\tilde{v}}^\top$ will explode), while those of $\boldsymbol{\Sigma}_{wr}$ remain bounded. We allow for a certain degree of cross-sectional dependence of $\boldsymbol{w}_{rt}$, as discussed later in the simulation study. The number of latent factors, K, is assumed to be known in this section.

We further assume that factors' loadings, $\boldsymbol{\beta}_{\tilde{v}}$, can partially explain expected returns,

$$\tilde{\boldsymbol{\mu}}_r = \boldsymbol{\mu}_r + \frac{1}{2}\boldsymbol{\Upsilon}_r = \boldsymbol{\beta}_{\tilde{v}}\boldsymbol{\lambda}_{\tilde{v}} + \boldsymbol{\alpha}, \tag{2}$$

where $\boldsymbol{\Upsilon}_r = \big(\text{var}(r_{1t}), \ldots, \text{var}(r_{Nt})\big)^\top$, $\boldsymbol{\lambda}_{\tilde{v}}$ denote risk premia associated with $\tilde{\boldsymbol{v}}_t$, and $\boldsymbol{\alpha}$ is a vector of pricing errors. The extra term $\frac{1}{2}\boldsymbol{\Upsilon}_r$ is added to the mean log excess returns in equation (2) due to the Jensen's inequality.[3] In addition, we assume that each asset's pricing error, α_i, is independently and identically distributed (IID) and cross-sectionally independent of factor loadings, with a zero mean and finite standard deviation. This form of model misspecification has been commonly used in the past literature (e.g., Kan et al. (2013), Gospodinov et al. (2014), Giglio and Xiu (2021), and Bryzgalova et al. (2023)) and has a clear economic interpretation. Equation (2) is equivalent to a log SDF that is linear in latent factors $\tilde{\boldsymbol{v}}_t$, described as follows:[4]

$$m_t = 1 - \boldsymbol{\lambda}_{\tilde{v}}^\top \tilde{\boldsymbol{v}}_t. \tag{3}$$

Since $\tilde{\boldsymbol{v}}_t$ have an identity covariance matrix, their risk prices are identical to risk premia.

We represent the covariance-stationary factor g_t as a sum of a moving average of asset return shocks plus other shocks, including measurement error, not spanned by financial markets:

$$g_t = \mu_g + \sum_{s=0}^{\bar{S}} \tilde{\rho}_s \underbrace{\tilde{\boldsymbol{\eta}}_g^\top \tilde{\boldsymbol{v}}_{t-s}}_{f_{t-s}} + w_{gt}, \quad \tilde{\boldsymbol{\eta}}_g^\top \tilde{\boldsymbol{\eta}}_g = 1, \tag{4}$$

[3] The approximation in equation (2) is exact under the lognormality assumption of asset returns.

[4] It is more appropriate to assume the log SDF as $m_t - \mathbb{E}_{t-1}[m_t] = \boldsymbol{\lambda}_{\tilde{v}}^\top \tilde{\boldsymbol{v}}_t$. However, since we study only excess returns, the (un)conditional mean of the SDF cannot be determined. For simplicity, we assume the function form in equation (3).

where μ_g is the unconditional mean of g_t, f_t is the spanned component that potentially drives both g_t and asset returns, $\{\tilde{\rho}_s\}_{s=0}^{\bar{S}}$ are square-summable, and w_{gt} is a potentially autocorrelated shock unrelated to $\tilde{\boldsymbol{v}}_t$ and $\boldsymbol{w}_{rt}$.

As long as the priced component of g_t is covariance stationary, the above representation must exist, possibly with $\bar{S} = \infty$ (by virtue of the Wold theorem), and requires $\tilde{\boldsymbol{\eta}}_g$ to be constant for all lags s. Furthermore, using a finite $\bar{S}$ must yield a finite approximation error relative to the true priced component, due to the square integrability property of the MA coefficients in the Wold representation. Note that since f_t is a white noise innovation, we can interpret $\{\tilde{\rho}_s\}_{s=0}^{\bar{S}}$ as g_t's impulse responses to the asset returns' shock f_t.[5] In other words,

$$\rho_s = \mathbb{E}\left[g_{t+s}|f_t = 1; g_{t-1}, \boldsymbol{r}_{t-1}\right] - \mathbb{E}\left[g_{t+s}|f_t = 0; g_{t-1}, \boldsymbol{r}_{t-1}\right], \tag{5}$$

and ρ_s is analogous to the local projection (LP, see, e.g., Jordà (2005)) coefficient of g_{t+s} on f_t, where the latter is identified leveraging a large cross-section of asset returns. That is, our framework, like LPs, recovers the impulse response of g to financial shocks avoiding the fragilities brought about by postulating and selecting a stringent autoregressive structure (Olea et al. (2024), Bryzgalova et al. (2024)). Furthermore, since we make use of the existence of a MA representation in the construction of $\mathbb{E}\left[g_{t+s}|f_t = 0; g_{t-1}, \boldsymbol{r}_{t-1}\right]$, we learn from the data about ρ_s not only from g_{t+s} but also from all other leads and lags of g_t, arguably gaining efficiency. Moreover, our approach controls directly for the MA structure implied by a sequence of linear projections[6] and uses only one estimating equation (equation (4), instead of $\bar{S}$ linear regressions with correlated residuals), hence greatly simplifying inference.

Several features of equation (4) are noteworthy. First, in theory, $\bar{S}$ can be $+\infty$, but we truncate the number of lags to ensure realistic estimation in finite samples. Second, g_t can react to both current and lagged asset return shocks $\tilde{\boldsymbol{v}}_t$. This assumption is motivated by the fact that asset prices are jump variables: news about current and future economic states are immediately incorporated into asset prices, whereas nontradable factors might respond to the same news with delay. The slow responses of nontraded economic variables to financial market shocks are also related to past literature showing that asset returns can predict macro variables (e.g., Liew and Vassalou (2000), Ang et al. (2006), and Bryzgalova et al. (2024)). Third, when

[5]We do not interpret f_t as a structural shock, so the impulse responses of g_t to f_t purely quantify the lead-lag correlations rather than the causal relationship between asset returns and g_t.

[6]Lusompa (2023) shows that the autocorrelation process of LP can be written as a Moving Average process of the Wold errors and impulse responses, and accounting for this dependency leads to more efficient estimates.

g_t correlates with only the contemporaneous asset return shocks (i.e., $\tilde{\rho}_s = 0$ for $s > 0$), the model reduces to the setting studied in Giglio and Xiu (2021).[7]

We now use several examples to illustrate how the general framework in Equations (1)–(4) maps into canonical representative agent models imposing particular parametric restrictions.

Example 1. *Adrian et al. (2014) measure a financial intermediary SDF, i.e., $m_t = 1 - \lambda \cdot LevFac_t$, where $LevFac_t$ is the shock to the leverage of security broker-dealers. To map our framework into theirs, we impose in equations (1)–(4) that $\tilde{v}_t = f_t = LevFac_t$, $\bar{S} = 0$, $\tilde{\rho}_0 = 1$, and g_t is a noisy proxy for $LevFac_t$ with a measurement error w_{gt}.*

Example 2. *In the canonical long-run consumption risk model of Bansal and Yaron (2004), the log consumption growth is modeled as $\Delta c_t = x_{t-1} + \sigma_{t-1}\eta_t$, where σ_{t-1} is the stochastic volatility process, x_{t-1} is the conditional consumption mean following an AR(1) process, $x_t = \rho_x x_{t-1} + \varphi_e \sigma_{t-1} e_t = \sum_{s=0}^{\infty} \varphi_e \rho_x^s \sigma_{t-s-1} e_{t-s}$, and $\sigma_{t-1}\eta_t$ is the short-run consumption shock. Within this framework, the log SDF is linear in three independent shocks, i.e., $m_t - \mathbb{E}_{t-1}(m_t) = \lambda_{m,\eta}\sigma_{t-1}\eta_t - \lambda_{m,e}\sigma_{t-1}e_t - \lambda_{m,\omega}\sigma_\omega \omega_t$ (ω_t is the shock to σ_t^2. See Internet Appendix IA.4 for details).*

The SDF in equation (3) maps into the Bansal and Yaron (2004) one imposing the following restrictions: i) $\tilde{\boldsymbol{v}}_t = (\sigma_{t-1}\eta_t, \sigma_{t-1}e_t, \sigma_\omega\omega_t)^\top$ and ii) $\boldsymbol{\lambda}_{\tilde{v}} = (-\lambda_{m,\eta}, \lambda_{m,e}, \lambda_{m,\omega})^\top$. To estimate the risk premium of the short-run consumption shock, $\sigma_{t-1}\eta_t$, we need to impose in equation (4) that $\bar{S} = 0$, $\tilde{\rho}_0 = 1$, $f_t = \sigma_{t-1}\eta_t$, and $w_{gt} = x_{t-1}$. Furthermore, to identify the risk premia of the shock to the conditional consumption mean, we need a different set of restrictions in equation (4): $\bar{S} = \infty$, $\rho_0 = 0$, $\rho_s = \varphi_e \rho_x^{s-1}$ for $s \geq 1$, $f_t = \sigma_{t-1}e_t$, and $w_{gt} = \sigma_{t-1}\eta_t$. In the empirical application, we consider the estimation for both $\bar{S} = 0$ and $\bar{S} >> 0$ to capture the risk premia of both short-run and long-run consumption shocks.

We next define the risk premium of g_t by extending the approach of Giglio and Xiu (2021). In their framework, g_t's risk premium is defined as the negative of the covariance between g_t and the SDF, $\lambda_g = -\text{cov}(g_t, m_t)$.[8] When g_t is a traded log excess return, the fundamental asset pricing equation, $\mathbb{E}\big[\exp(m_t + g_t + r_{ft})\big] = 1$, implies $\mathbb{E}[g_t] + \frac{1}{2}\text{var}(g_t) = -\text{cov}(g_t, m_t)$ under the joint log normality assumption. For a nontradable factor, one can interpret $-\text{cov}(g_t, m_t)$ as the pseudo expected excess return of g_t as if it were tradable. In other words, $-\text{cov}(g_t, m_t)$ is the

[7] Since their paper uses original rather than log returns, this statement is precise with the exception of the log-linearization approximation error.

[8] This definition is consistent with Cochrane (2009, Chapter 6).

risk premium on an asset that delivers a payoff that grows at the rate of g_t. We expand their definition by allowing for an entire term structure of risk premia. Specifically, the (average per-period) risk premium of g from $t-1$ to $t+S$ ($0 \leq S \leq \bar{S}$) is defined as the multiperiod covariance between the factor and the SDF, divided by the number of holding periods, as follows:

$$\lambda_g^S = -\frac{\text{cov}(m_{t-1\to t+S}, g_{t-1\to t+S})}{1+S} = \frac{\sum_{\tau=0}^{S}\sum_{s=0}^{\tau}\tilde{\rho}_s}{1+S} \cdot \underbrace{\tilde{\boldsymbol{\eta}}_g^\top \boldsymbol{\lambda}_{\tilde{\boldsymbol{v}}}}_{\lambda_f}. \tag{6}$$

There are two ways to interpret the definition in equation (6). First, λ_f is the risk premium of the spanned component ($f_t = \tilde{\boldsymbol{\eta}}_g^\top \tilde{\boldsymbol{v}}_t$) driving both asset returns and g_t, and $\frac{\sum_{\tau=0}^{S}\sum_{s=0}^{\tau}\tilde{\rho}_s}{1+S}$ is the per-period loading of $g_{t-1\to t+S}$ on multiperiod asset return shocks $f_{t-1\to t+S}$. Hence, λ_g^S, the risk premium of g over an investment horizon of $(1+S)$ periods, equals its per-period loadings on f multiplied by f's risk premium.

Second, as established below, we can interpret λ_g^S as the risk premium of the *horizon-specific* mimicking portfolio hedging against $g_{t-1\to t+S}$, with portfolio weights $\boldsymbol{w}^{MP}$ as follows:

$$\boldsymbol{w}^{MP} = \text{cov}(\boldsymbol{r}_{t-1\to t+S})^{-1}\text{cov}(\boldsymbol{r}_{t-1\to t+S}, g_{t-1\to t+S}).$$

The risk premium of this portfolio, normalized by the number of holding periods, is

$$\lambda_g^{MP} = \frac{\left(\mathbb{E}[\boldsymbol{r}_{t-1\to t+S}] + \frac{1}{2}\boldsymbol{\Upsilon}(\boldsymbol{r}_{t-1\to t+S})\right)^\top \boldsymbol{w}^{MP}}{1+S} = \left(\mathbb{E}[\boldsymbol{r}_t] + \frac{1}{2}\boldsymbol{\Upsilon}_r\right)^\top \text{cov}(\boldsymbol{r}_t)^{-1}\boldsymbol{\beta}_{\tilde{v}}\frac{\text{cov}(\tilde{\boldsymbol{v}}_{t-1\to t+S}, g_{t-1\to t+S})}{1+S},$$

where the last equality uses the assumption that $\boldsymbol{v}_t$ are serially uncorrelated and $\boldsymbol{w}_{r,t-1\to t+S}$ are orthogonal to g. We relax the assumption of uncorrelated $\tilde{\boldsymbol{v}}_t$ in Section 2.2.

Using Proposition A.1 in the Appendix, we simplify the risk premium of g_t's mimicking portfolio and show that, as the number of test assets goes to infinity, $\lambda_g^{MP} \to \frac{\boldsymbol{\lambda}_{\tilde{v}}^\top \text{cov}(\tilde{\boldsymbol{v}}_{t-1\to t+S}, g_{t-1\to t+S})}{1+S} = -\frac{\text{cov}(m_{t-1\to t+S}, g_{t-1\to t+S})}{1+S}$, where $m_{t-1\to t+S} = \sum_{\tau=0}^{S} m_{t+\tau-1,t+\tau} = 1 + S - \boldsymbol{\lambda}_{\tilde{v}}^\top \tilde{\boldsymbol{v}}_{t-1\to t+S}$. Therefore, our definition of g_t's risk premium in equation (6) is asymptotically equivalent to the risk premium of the horizon-specific mimicking portfolio in a large cross-section.

Example 3. *Suppose that the IID CAPM holds: $f_t = \tilde{v}_t = r_t^{mkt}$, with r_t^{mkt} independent over time and normalized to have unit variance. The SDF is then $m_t = 1 - \lambda_{mkt} r_t^{mkt}$. For any factor g_t that follows the process in equation (4), we can compute the term structure of its risk premia using the definition in equation (6), as follows:*

$$\lambda_g^S = -\frac{\text{cov}(m_{t-1\to t+S}, g_{t-1\to t+S})}{1+S} = \left[\beta_0^g + \frac{1}{1+S} \overbrace{\sum_{\tau=1}^{S}\sum_{s=1}^{\tau}\beta_s^g}^{\text{“forward”-market-}\beta s}\right]\lambda_{mkt},$$

where the forward market betas, $\beta_S^g \equiv \frac{\text{cov}(g_{t-1+s\to t+s}, r_t^{mkt})}{\sigma_{mkt}^2}$*, captures the predictability of* g*. There are two takeaways from this example. First, the term structure of factor risk premia is determined by how the same priced asset return shocks are propagated through the tested factor. Second, the mimicking portfolio based on purely the single-period market beta (*β_0^g*) is generally uninformative about the multi-period risk premia, since it ignores the information embedded in forward betas.*

In the data, asset return factors, $\tilde{\boldsymbol{v}}_t$, are unidentified. That is, one can only estimate a linear rotation of $\tilde{\boldsymbol{v}}_t$, denoted by $\boldsymbol{v}_t = \boldsymbol{H}\tilde{\boldsymbol{v}}_t$, where $\boldsymbol{H}$ is a $K \times K$ nonsingular matrix. Since $\tilde{\boldsymbol{v}}_t \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}_K, \boldsymbol{I}_K)$, we have that $\boldsymbol{\Sigma}_v \equiv \text{cov}(\boldsymbol{v}_t) = \boldsymbol{H}\boldsymbol{H}^\top$. Even though $\tilde{\boldsymbol{v}}_t$ cannot be identified, g_t's risk premium is well-defined. In particular, the identification of λ_g^S builds upon the rotation invariance property emphasized in Giglio and Xiu (2021). The rotation invariance can be easily seen by rewriting the model as follows:

$$\begin{aligned}
&\boldsymbol{r}_t = \boldsymbol{\alpha} + \underbrace{\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\boldsymbol{\lambda}_{\tilde{v}}}_{\boldsymbol{\lambda}_v} - \frac{1}{2}\boldsymbol{\Upsilon}_r + \underbrace{\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\tilde{\boldsymbol{v}}_t}_{\boldsymbol{v}_t} + \boldsymbol{w}_{rt}, \quad g_t = \mu_g + \sum_{s=0}^{\bar{S}}\tilde{\rho}_s\underbrace{\tilde{\boldsymbol{\eta}}_g^\top\boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_g^\top}\underbrace{\boldsymbol{H}\tilde{\boldsymbol{v}}_{t-s}}_{\boldsymbol{v}_{t-s}} + w_{gt}, \text{ and} \\
&m_t = 1 - \boldsymbol{\lambda}_v^\top(\boldsymbol{H}^{-1})^\top\boldsymbol{H}^{-1}\boldsymbol{v}_t = 1 - \boldsymbol{\lambda}_v^\top\boldsymbol{\Sigma}_v^{-1}\boldsymbol{v}_t, \quad \lambda_g^S = \frac{\sum_{\tau=0}^{S}\sum_{s=0}^{\tau}\tilde{\rho}_s}{1+S}\cdot\underbrace{\tilde{\boldsymbol{\eta}}_g^\top\boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_g}\underbrace{\boldsymbol{H}\boldsymbol{\lambda}_{\tilde{v}}}_{\boldsymbol{\lambda}_v}.
\end{aligned} \tag{7}$$

Therefore, the most important quantity in our paper, λ_g^S, is well identified.

Remark 1. λ_f *in equation (6) can be interpreted as the risk price of* f_t *after controlling for the omitted sources of priced risk in the SDF. Let* $\boldsymbol{v}_t = \boldsymbol{H}\tilde{\boldsymbol{v}}_t$*, where* $\boldsymbol{H}^\top = (\tilde{\boldsymbol{\eta}}_g, \boldsymbol{H}_1)$ *is a* $K \times K$ *nonsingular matrix, and* $\boldsymbol{H}_1^\top\tilde{\boldsymbol{\eta}}_g = \boldsymbol{0}$*. Under this formulation,* $\boldsymbol{v}_t = (f_t, \boldsymbol{u}_t^\top)^\top$*,* $\boldsymbol{u}_t = \boldsymbol{H}_1^\top\tilde{\boldsymbol{v}}_t$*, and* $f_t \perp \boldsymbol{u}_t$*. The log SDF in equation (3) can be rewritten as* $m_t = 1 - \boldsymbol{\lambda}_v^\top\boldsymbol{\Sigma}_v^{-1}\boldsymbol{v}_t = 1 - \lambda_f f_t - \boldsymbol{\lambda}_u^\top(\boldsymbol{H}_1^\top\boldsymbol{H}_1)^{-1}\boldsymbol{u}_t$*, where* $\lambda_f = \tilde{\boldsymbol{\eta}}_g^\top\boldsymbol{\lambda}_{\tilde{v}}$ *and* $\boldsymbol{\lambda}_u = \boldsymbol{H}_1^\top\boldsymbol{\lambda}_{\tilde{v}}$*. Using this particular SDF representation, we can decompose the variance of the SDF, which is equivalent to the squared maximal Sharpe ratio in the economy, as* $\text{var}(m_t) = \lambda_f^2 + \text{var}(\boldsymbol{\lambda}_u^\top(\boldsymbol{H}_1^\top\boldsymbol{H}_1)^{-1}\boldsymbol{u}_t)$*. Hence,* λ_f *can be interpreted as the (per period) model-implied Sharpe ratio of the* f_t *shock;* $\lambda_f^2/\text{var}(m_t)$ *quantifies the relative importance of* f_t *in the SDF, conditioned that* f_t *is given the largest power in the log SDF to explain the cross-section of average returns. If* f_t *is strongly identified in a macro factor, we* *can examine the largest role that this macro state variable plays in the SDF.*

Estimating the confidence bands – or better, the statistical uncertainty – of λ_g^S is challenging in the frequentist framework. Specifically, λ_g^S is a function of $\boldsymbol{\rho}_g$, $\boldsymbol{\eta}_g$, and $\boldsymbol{\lambda}_v$, where the first two parameters depend on each other. Hence, the frequentist asymptotic covariance matrix of λ_g^S is quite complex despite its closed-form expression outlined above. Consequently, we adopt a Bayesian framework to provide valid inference for all model parameters and present it in the next subsection.

2.1 Bayesian Estimation of Risk Premia

This subsection describes our hierarchical Bayesian framework. We first consider the *time series* dimension, which is needed to estimate the joint posterior distribution of asset returns' latent factors and their loadings, expected asset returns, g_t's loadings on the latent factors, and the precision matrices of error terms. We make the following distributional assumptions:

$$g_t = \mu_g + \sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top (\boldsymbol{v}_{t-s} - \boldsymbol{\mu}_{\boldsymbol{v}}) + w_{gt}, \quad w_{gt} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{wg}^2), \quad \boldsymbol{v}_t \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{\mu}_{\boldsymbol{v}}, \boldsymbol{\Sigma}_{\boldsymbol{v}}), \tag{8}$$

$$\boldsymbol{r}_t = \boldsymbol{\mu}_r + \boldsymbol{\beta}_v(\boldsymbol{v}_t - \boldsymbol{\mu}_v) + \boldsymbol{w}_{rt}, \quad \boldsymbol{w}_{rt} \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}_N, \boldsymbol{\Sigma}_{wr}), \quad \boldsymbol{\Sigma}_{wr} = \text{diag}\{\sigma_{1,wr}^2, \dots, \sigma_{N,wr}^2\}, \text{ and} \tag{9}$$

$$\boldsymbol{v}_t \perp w_{gt} \perp \boldsymbol{w}_{rt}, \text{ and let } \boldsymbol{\rho}_g = (\mu_g, \rho_0, \dots, \rho_{\bar{S}})^\top, \tag{10}$$

where $\boldsymbol{v}_t$ are linear and nonsingular rotations of the true K latent factors $\tilde{\boldsymbol{v}}_t$. Since these rotations are arbitrary, we need to estimate their unconditional means ($\boldsymbol{\mu}_{\boldsymbol{v}}$) and covariance matrix ($\boldsymbol{\Sigma}_{\boldsymbol{v}}$). Direct modeling of $\boldsymbol{\mu}_v$ is critical for obtaining a proper posterior distribution of expected excess returns $\boldsymbol{\mu}_r$.[9] According to equation (10), the error terms, w_{gt} and $\boldsymbol{w}_{rt}$, are orthogonal, which implies that we can estimate the model parameters in g_t and $\boldsymbol{r}_t$ separately.

The systems in (8) and (9) introduce a potential degree of misspecification relative to the true data-generating processes described in equations (2) and (4). First, the error w_{gt} could be serially correlated. As Müller (2013) shows, posteriors are still asymptotically normal and centered at the maximum likelihood estimate under this assumption, although the canonical posterior covariance matrix of the model parameters is incorrect and should be replaced with a sandwich covariance matrix. We incorporate this correction within our method.

[9] The sample average of $\boldsymbol{r}_t$ is $\boldsymbol{\mu}_r + \boldsymbol{\beta}_v \frac{1}{T}\sum_{t=1}^T(\boldsymbol{v}_t - \boldsymbol{\mu}_v) + \frac{1}{T}\sum_{t=1}^T \boldsymbol{w}_{rt}$. If we always demean the latent factors to have zero sample averages, the first source of uncertainty about $\boldsymbol{\mu}_r$, originated from $\frac{1}{T}\sum_{t=1}^T(\boldsymbol{v}_t - \boldsymbol{\mu}_v)$, will disappear. Consequently, the credible intervals for $\boldsymbol{\mu}_r$ will be too tight if we do not directly model $\boldsymbol{\mu}_v$.

Second, $\boldsymbol{\Sigma_{wr}}$ is assumed to be diagonal. Our posterior characterization below does not require this assumption, and indeed, we impose it only to avoid numerical problems when considering very large cross-sectional dimensions (i.e., when the number of assets approaches or exceeds the time series dimension). However, as we will show through simulations, the diagonal assumption does not have material effects on the posterior distributions. Hence, this assumption is harmless. This robustness result is not surprising since, in a frequentist setting, this type of misspecification would affect only efficiency but not consistency.

We assign the standard uninformative prior distributions to the time series parameters

$$\begin{aligned}&\pi(\boldsymbol{\rho}_g, \boldsymbol{\eta}_g, \sigma^2_{wg}) \propto (\sigma^2_{wg})^{-1}, \quad \pi(\boldsymbol{v}) \propto 1, \quad \pi(\boldsymbol{\mu}_v, \boldsymbol{\Sigma}_v) \propto |\boldsymbol{\Sigma}_v|^{-\frac{K+1}{2}}, \text{ and} \\ &\pi(\boldsymbol{\beta}_v) \propto 1, \quad \pi(\boldsymbol{\mu}_r, \boldsymbol{\Sigma}_{wr}) \propto |\boldsymbol{\Sigma}_{wr}|^{-\frac{N+1}{2}}.\end{aligned} \tag{11}$$

In the *cross-sectional* dimension, conditional on the recovered sources of risk v_t in the time series dimension, the SDF and its risk prices, λ_v, can then be recovered using the Bayesian-SDF estimator (B-SDF) in Definition 1 of Bryzgalova et al. (2023). That is, conditional on the recovered v_t being the sources of risk driving the cross-section, we have the SDF

$$m_t = 1 - \boldsymbol{\lambda}_v^\top \boldsymbol{\Sigma}_v^{-1} \boldsymbol{v}_t \;\Rightarrow\; \tilde{\boldsymbol{\mu}}_r = \boldsymbol{\beta}_v \boldsymbol{\lambda}_v. \text{[10]} \tag{12}$$

Recall that we nevertheless allow for pricing errors as outlined in (2). With extensive simulation studies, we show in Section 3 that this approach delivers valid posterior distributions.

Within the frequentist paradigm, constructing proper inference for the system in equations (8)–(12) is, if not infeasible, at least a daunting task. As we are about to show in Proposition 1 below, this is both simple and transparent within the Bayesian paradigm.

There are two reasons for this. First, a joint distribution, say $p(x, y)$, can be traced by generating a Markov chain that sequentially samples from $p(x|y)$ and $p(y|x)$ – the so-called Gibbs sampling.

Second, the hierarchical structure of the time series and cross-sectional layers of the estimation problem yields well-defined and well-understood conditional posterior distributions. Specifically, if $\boldsymbol{v}_t$ were known (i.e., conditioning on it), equation (8) would simply be an ordinary linear regression problem with well-known properties: in a Bayesian setting, under diffuse and/or conjugate priors, a normal-inverse-gamma posterior distribution (i.e., the analogue of the t-distribution that would arise for frequentist inference in this case).

[10] $\tilde{\boldsymbol{\mu}}_r = \boldsymbol{\mu}_r + \frac{1}{2}\boldsymbol{\Upsilon}_r$, where both $\boldsymbol{\mu}_r$ and $\boldsymbol{\Upsilon}_r$ are estimated in the time series step.

Similarly, if $\boldsymbol{v}_t$ were known, equation (9) would simply be a canonical multivariate linear regression, thereby yielding (under diffuse and/or conjugate priors) a well-known posterior distribution: a normal-inverse-Wishart (the Bayesian analogue of the frequentist multivariate t-distribution result).

Furthermore, conditional on knowing both the parameters in equation (9) and the data, the distribution of the latent factors $\boldsymbol{v}_t$ can be obtained by inverting its relationship with asset returns. Finally, conditional on the parameters and latent factors in the time series layer, the distribution of the risk prices, $\boldsymbol{\lambda}_v$, simply follows from Definition 2 of Bryzgalova et al. (2023). Note that this layer is fundamental since it de facto selects which of (and how) the latent drivers $\boldsymbol{v}_t$ are actually sources of priced risk – the crucial stage for measuring the risk premia associated with g_t.

We formalize this hierarchal characterization of the posterior in the proposition below and derive it in Internet Appendix IA.1.1.

Proposition 1 (Gibbs sampler of the baseline model)**.** *Under the assumptions in equations (8)–(12), the posterior distribution of the model parameters can be sampled from the following conditional distributions:*

(1) Conditional on the data, $\{g_t\}_{t=1+\bar{S}}^T$, and latent factors, $\{\boldsymbol{v}_t\}_{t=1}^T$, the parameters of the g_t process (σ_{wg}^2, $\boldsymbol{\rho}_g$, and $\boldsymbol{\eta}_g$) follow the normal-inverse-gamma distribution in equations (IA.1)–(IA.3) of Internet Appendix IA.1.1. For point identification purposes, draws of $\boldsymbol{\rho}_g$ and $\boldsymbol{\eta}_g$ are normalized such that $\boldsymbol{\eta}_g^\top \boldsymbol{\eta}_g = 1$.

(2) Conditional on asset returns, $\{\boldsymbol{r}_t\}_{t=1}^T$, and latent factors, the parameters of the $\boldsymbol{r}_t$ process ($\boldsymbol{\Sigma}_{wr}$ and $\boldsymbol{B}_r^\top = (\boldsymbol{\mu}_r, \boldsymbol{\beta}_v)$) follow the normal-inverse-Wishart distribution in equations (IA.4)–(IA.5) of Internet Appendix IA.1.1.

(3) Conditional on asset returns and $(\boldsymbol{\mu}_r, \boldsymbol{\beta}_v, \boldsymbol{\Sigma}_{wr})$, the latent factors, $\boldsymbol{v}_t$, their mean, and covariance matrix can be sampled from

$$\boldsymbol{v}_t \mid \boldsymbol{r}_t, \boldsymbol{\mu}_r, \boldsymbol{\beta}_v, \boldsymbol{\Sigma}_{wr}, \boldsymbol{\mu}_v, \boldsymbol{\Sigma}_v \sim \mathcal{N}\left(\left(\boldsymbol{\beta}_{\boldsymbol{v}}^\top \boldsymbol{\Sigma}_{\boldsymbol{wr}}^{-1} \boldsymbol{\beta}_{\boldsymbol{v}}\right)^{-1} \left[\boldsymbol{\beta}_{\boldsymbol{v}}^\top \boldsymbol{\Sigma}_{\boldsymbol{wr}}^{-1} (\boldsymbol{r}_t - \boldsymbol{\mu}_r + \boldsymbol{\beta}_{\boldsymbol{v}} \boldsymbol{\mu}_v)\right],\ \left(\boldsymbol{\beta}_{\boldsymbol{v}}^\top \boldsymbol{\Sigma}_{\boldsymbol{wr}}^{-1} \boldsymbol{\beta}_{\boldsymbol{v}}\right)^{-1} \right), \quad (13)$$

$$\boldsymbol{\Sigma}_v \mid \{\boldsymbol{v}_t\}_{t=1}^T \sim \mathcal{W}^{-1}\Bigl(T-1,\ \sum_{t=1}^T (\boldsymbol{v}_t - \bar{\boldsymbol{v}})(\boldsymbol{v}_t - \bar{\boldsymbol{v}})^\top\Bigr),\ \textit{and} \quad (14)$$

$$\boldsymbol{\mu}_v \mid \boldsymbol{\Sigma}_v, \{\boldsymbol{v}_t\}_{t=1}^T \sim \mathcal{N}\left(\bar{\boldsymbol{v}},\ \boldsymbol{\Sigma}_v/T\right), \tag{15}$$

where $\mathcal{N}(\cdot)$ and $\mathcal{W}^{-1}(\cdot)$ denote, respectively, the normal and inverse-Wishart distributions.

(4) Conditional on the posterior draws from the time series steps (1)–(3), the posterior distribution of $\boldsymbol{\lambda}_v$ is a Dirac distribution at $(\boldsymbol{\beta}_v^\top \boldsymbol{\beta}_v)^{-1}\boldsymbol{\beta}_v^\top \tilde{\boldsymbol{\mu}}_r$, yielding a Dirac conditional posterior for the term structure of g_t's risk premia at $\lambda_g^S = \frac{\sum_{\tau=0}^{S}\sum_{s=0}^{\tau}\rho_s}{1+S} \cdot \boldsymbol{\eta}_g^\top \boldsymbol{\lambda}_{\boldsymbol{v}}$, where $0 \le S \le \bar{S}$.

Several features of our Bayesian Gibbs sampler are noteworthy. First, although we do not know in closed-form the joint distribution of all parameters, all conditional distributions, such as inverse-gamma, multivariate normal, and inverse-Wishart distributions, are well-defined and standard.

Second, we follow Müller (2013) and adjust the posterior covariance matrix of $\boldsymbol{\rho}_g$ and $\boldsymbol{\eta}_g$ for the autocorrelation in the residuals, w_{gt} and $\boldsymbol{w}_{rt}$, using the Newey and West (1987) type of sandwich estimator.[11]

Third, the posterior distribution of $\boldsymbol{v}_t$ in Step 3 of Proposition 1 ignores the information embedded in g_t, balancing the trade-off between model simplicity and estimation efficiency. Since g_t depends on many lags of the latent factors, incorporating its information in estimating $\boldsymbol{v}_t$ is feasible but requires a more computationally demanding approach, such as the Kalman filter. Since we consider large cross-sections of test assets, the discarded information is negligible as $N \to \infty$. Finally, in empirical applications, not conditioning on g_t in the extraction of $\boldsymbol{v}_t$ provides a level playing field when comparing the estimated risk premia of different g_t.

Fourth, Proposition 1 does *not* require a diagonal $\boldsymbol{\Sigma}_{wr}$. Nevertheless, for empirical applications where N is close to the time series sample size, we impose diagonality to avoid numerical difficulties. Our simulation studies confirm that the assumption of a diagonal $\boldsymbol{\Sigma}_{wr}$ does not result in invalid confidence intervals, even though $\boldsymbol{w}_{rt}$ is cross-sectionally correlated in the hypothetical true data-generating process. In contrast, in empirical applications where the number of test assets is relatively small (i.e., $N \le 50$, such as in the cross-section of corporate bonds), we use a nondiagonal $\boldsymbol{\Sigma}_{wr}$ in estimation.

Fifth, the cross-sectional dimension (Step 4 in Proposition 1) defines latent factors' risk premia as $(\boldsymbol{\beta}_v^\top \boldsymbol{\beta}_v)^{-1}\boldsymbol{\beta}_v^\top \tilde{\boldsymbol{\mu}}_r$ and, via the sequential resampling, accounts for the uncertainty

[11] The number of lags is set to be $\bar{S}$ since $w_{gt}\boldsymbol{x}_t$ and $w_{g,t-l}\boldsymbol{x}_{t-l}$ become serially uncorrelated for $l > \bar{S}$, where x_t denote the regressors in g_t's equation and is the linear transformation of latent factors $\{v_{t-s}\}_{s=0}^{\bar{S}}$.

about the expected returns, the factor loadings, and the latent factors' means $\boldsymbol{\mu}_v$.

In addition to risk premia estimates, our Bayesian framework can produce valid posterior distributions for other economic quantities of interest, including, but not limited to, the time series fit in g_t's equation (R_g^2), cumulative impulse responses of g_t to the asset return shocks ($\{\tilde{\rho}_s\}_{s=0}^{\bar{S}}$), and the cross-sectional fit in explaining average returns.

Past literature often adopts the Fama-MacBeth regression to estimate factors' risk premia. In Proposition 1, steps 2–4 echo the time series and cross-sectional steps of the Bayesian Fama-MacBeth in Bryzgalova et al. (2023) for principal components of asset returns. Step 1 is the additional step that models the joint dynamics of asset returns and g_t. As Giglio and Xiu (2021) argue, estimating factors' risk premia using principal components of asset returns can avoid the omitted variable bias and attenuation bias from measurement errors.

Finally, the traditional Fama-MacBeth regression suffers from weak identification (see, e.g., Kan and Zhang (1999a,b)), particularly for macro factors. One contribution of our paper is to use the factors' cumulative loadings on asset returns, proxied by $\{\tilde{\rho}_s\}_{s=0}^{\bar{S}}$, to identify their risk premia. In short, we will show in both simulation studies and real-world data that our Bayesian estimates are not only robust to the weak identification but, more importantly, help recover the risk premia of persistent macro factors.

2.2 Time-Varying Risk Premia and Their Term Structures

From an economic standpoint, a salient feature of macro-finance equilibrium models is the time variation in risk premia. In this section, we extend our Bayesian framework for estimating time-varying term structures.

We now require the SDF to price assets *conditionally*; that is,[12]

$$\underbrace{\mathbb{E}_t[r_{i,t+1}] + \frac{1}{2}\text{var}_t(r_{i,t+1})}_{\tilde{\mu}_{r,i,t}} = -\text{cov}_t(m_{t+1}, r_{i,t+1}), \quad i = 1 \ldots N, \tag{16}$$

where $\mathbb{E}_t$ denotes the conditional expectation at time t, and $\text{var}_t(r_{i,t+1})$ is the conditional variance of $r_{i,t+1}$. Throughout our paper, we consider homoskedastic asset returns; hence, $\text{var}_t(r_{i,t+1})$ is constant over time. We define $\boldsymbol{\Upsilon}_r$ as $\big(\text{var}_t(r_{1,t+1}), \ldots, \text{var}_t(r_{N,t+1})\big)^\top$. Leveraging

[12] Since the SDF prices the log excess returns, we have $\mathbb{E}_t\big[\exp(m_{t+1} + r_{i,t+1} + r_{f,t+1})\big] = 1, \quad i = 1 \ldots N$. Hence, under the joint log normality assumption, we obtain equation (16).

Hansen and Jagannathan (1991), we focus on the conditional SDF projections on the space of returns as follows:

$$m_{t+1} = 1 - \boldsymbol{b}_t^\top \left(\boldsymbol{r}_{t+1} - \mathbb{E}_t[\boldsymbol{r}_{t+1}]\right), \quad \text{where} \quad \boldsymbol{b}_t = \text{cov}_t(\boldsymbol{r}_{t+1})^{-1} \tilde{\boldsymbol{\mu}}_{rt}. \tag{17}$$

The return process, as before, follows an approximate factor structure,

$$\boldsymbol{r}_t = \boldsymbol{\mu}_r + \boldsymbol{\beta}_{\tilde{v}} \tilde{\boldsymbol{v}}_t + \boldsymbol{w}_{rt}, \quad \tilde{\boldsymbol{v}}_t \perp \boldsymbol{w}_{rt}, \quad \mathbb{E}_{t-1}[\boldsymbol{w}_{rt}] = \boldsymbol{0}_N, \quad \mathbb{E}[\tilde{\boldsymbol{v}}_t] = \boldsymbol{0}_K, \tag{18}$$

where, importantly, the priced systematic factors $\tilde{\boldsymbol{v}}_t$ are potentially predictable. That is, $\tilde{\boldsymbol{v}}_t = \boldsymbol{\mu}_{\tilde{v},t-1} + \boldsymbol{\epsilon}_{\tilde{v}t}$, where $\boldsymbol{\mu}_{\tilde{v},t-1} \equiv \mathbb{E}_{t-1}[\tilde{\boldsymbol{v}}_t]$; hence $\boldsymbol{\mu}_{\tilde{v},t-1} \perp \boldsymbol{\epsilon}_{\tilde{v}t}$. We normalize the innovations to the latent factors such that $\text{cov}(\boldsymbol{\epsilon}_{\tilde{v}t}) = \boldsymbol{I}_K$.

As previously, unconditional mean returns are partially explained by $\boldsymbol{\beta}_{\tilde{v}}$ in equation (2). The only additional assumption that we require is that the eigenvalues of $\text{cov}(\boldsymbol{\mu}_{\tilde{v},t-1})$ are bounded. This formulation yields the SDF[13]

$$m_{t+1} = 1 - \boldsymbol{\lambda}_{\tilde{v}}^\top \boldsymbol{\epsilon}_{\tilde{v},t+1} - \boldsymbol{\mu}_{\tilde{v}t}^\top \boldsymbol{\epsilon}_{\tilde{v},t+1}, \tag{19}$$

where $\boldsymbol{\mu}_{\tilde{v}t}^\top \boldsymbol{\epsilon}_{\tilde{v},t+1}$ captures time-varying risk premia of asset return shocks.

Since the Wold representation requires the MA formulation to depend only on innovations, the process for g is modified as follows:

$$g_t = \mu_g + \sum_{s=0}^{\bar{S}} \tilde{\rho}_s \underbrace{\tilde{\boldsymbol{\eta}}_g^\top \boldsymbol{\epsilon}_{\tilde{v},t-s}}_{f_{t-s}} + w_{gt}, \quad \tilde{\boldsymbol{\eta}}_g^\top \tilde{\boldsymbol{\eta}}_g = 1. \tag{20}$$

That is, g is potentially driven by the innovations of the priced systematic factors $\tilde{\boldsymbol{v}}_t$. Hence, defining the *conditional* risk premia analogously as the unconditional ones, we have that the time-varying term structure of risk premia is given by

$$\lambda_{g,t-1}^S = -\frac{\text{cov}_{t-1}(m_{t-1\to t+S}, g_{t-1\to t+S})}{1+S} = \sum_{\tau=0}^{S} \sum_{s=0}^{\tau} \frac{\tilde{\rho}_s \tilde{\boldsymbol{\eta}}_g^\top \left(\boldsymbol{\lambda}_{\tilde{v}} + \mathbb{E}_{t-1}\left[\boldsymbol{\mu}_{\tilde{v},t+\tau-s-1}\right]\right)}{1+S}. \tag{21}$$

[13]Using equations (17) and (18), we can show that $\boldsymbol{b}_t = \left(\boldsymbol{\beta}_{\tilde{v}} \boldsymbol{\beta}_{\tilde{v}}^\top + \boldsymbol{\Sigma}_{wr}\right)^{-1} \left(\boldsymbol{\alpha} + \boldsymbol{\beta}_{\tilde{v}} \boldsymbol{\lambda}_{\tilde{v}} + \boldsymbol{\beta}_{\tilde{v}} \boldsymbol{\mu}_{\tilde{v},t}\right)$ and $\boldsymbol{r}_{t+1} - \mathbb{E}_t[\boldsymbol{r}_{t+1}] = \boldsymbol{\beta}_{\tilde{v}} \boldsymbol{\epsilon}_{\tilde{v},t+1} + \boldsymbol{w}_{r,t+1}$. Ignoring the unpriced shocks $\boldsymbol{w}_r$, we can represent the linear SDF as $m_{t+1} = 1 - \boldsymbol{\alpha}^\top \left(\boldsymbol{\beta}_{\tilde{v}} \boldsymbol{\beta}_{\tilde{v}}^\top + \boldsymbol{\Sigma}_{wr}\right)^{-1} \boldsymbol{\beta}_{\tilde{v}} \boldsymbol{\epsilon}_{\tilde{v},t+1} - (\boldsymbol{\lambda}_{\tilde{v}} + \boldsymbol{\mu}_{\tilde{v},t})^\top \boldsymbol{\beta}_{\tilde{v}}^\top \left(\boldsymbol{\beta}_{\tilde{v}} \boldsymbol{\beta}_{\tilde{v}}^\top + \boldsymbol{\Sigma}_{wr}\right)^{-1} \boldsymbol{\beta}_{\tilde{v}} \boldsymbol{\epsilon}_{\tilde{v},t+1}$. Following similar derivations as in Appendix A.1, we can derive that $m_{t+1} \to 1 - (\boldsymbol{\lambda}_{\tilde{v}} + \boldsymbol{\mu}_{\tilde{v},t})^\top \boldsymbol{\epsilon}_{\tilde{v},t+1}$ as $N \to \infty$.

Four important observations are in order. First, the dynamics of the conditional mean of the systematic risks, $\boldsymbol{\mu}_{\tilde{v},t-1}$, drive the time variation of the term structure of risk premia. Second, since by construction $\mathbb{E}[\boldsymbol{\mu}_{\tilde{v},t-1}] = 0$, the implied unconditional term structure is the same as that of equation (6), which was obtained with uncorrelated sources of systematic risk. That is, the estimator derived in Section 2.1 is consistent even in the presence of time-varying risk premia. Third, despite the added generality, the risk premia of g remain point-identified due to the rotation invariance property of our setting (See Appendix A.2). Fourth, to elicit the time variation of the term structure, we need to explicitly model the conditional mean process, that is, the dynamics of $\tilde{\boldsymbol{v}}$.

We assume that $\tilde{\boldsymbol{v}}_t$ are driven by some predictors, such as $\tilde{\boldsymbol{v}}_t$'s lags and p external variables $\boldsymbol{z}_t$. Let $\boldsymbol{x}_t = (\tilde{\boldsymbol{v}}_t^\top, \boldsymbol{z}_t^\top)^\top$, which follows a vector autoregressive (VAR) model of order q:[14]

$$\boldsymbol{x}_t = \boldsymbol{\phi}_0 + \boldsymbol{\phi}_1 \boldsymbol{x}_{t-1} + \cdots + \boldsymbol{\phi}_q \boldsymbol{x}_{t-q} + \boldsymbol{\epsilon}_{xt}, \quad \boldsymbol{\epsilon}_{xt} \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}_{K+p}, \boldsymbol{\Sigma}_{\epsilon x}). \tag{22}$$

The additional layer in equation (22) requires a minimal change to our Gibbs sampler to characterize the posterior distribution. The only deviation from Section 2.1 is that $\boldsymbol{v}_t$ follows a VAR process rather than an IID normal distribution. In particular, using the canonical diffuse prior $\pi(\boldsymbol{\phi}_0, \ldots, \boldsymbol{\phi}_q, \boldsymbol{\Sigma}_{\epsilon x}) \propto |\boldsymbol{\Sigma}_{\epsilon x}|^{-\frac{K+p+1}{2}}$, the conditional posterior of the parameters in this additional layer follows the usual normal-inverse-Wishart distribution and can be sampled accordingly. We summarize the Gibbs sampler in Proposition A2 of Appendix A.2 and derive it in Internet Appendix IA.1.2.

The time-varying framework in this subsection is closely connected to the literature on affine term structure models, such as Kim and Wright (2005) and Cochrane and Piazzesi (2008). For instance, in Cochrane and Piazzesi (2008), $\boldsymbol{x}_t$ in equation (22) contains three latent factors (level, slope, and curvature) of government bond yields, plus the bond-return forecasting factor in Cochrane and Piazzesi (2005). Besides, they also assume that risk prices of the shocks to latent factors are linear functions of the lagged bond-return forecasting factor. Unlike their paper, since we focus on estimating only risk premia, we do not need to model the dynamics of the risk-free rates. Hence, we always normalize m_t to have a constant mean.

Our paper shares some common modelling choices with Kelly et al. (2023) in that we study the log SDF linear in latent factors of equity excess returns, impose log normality, and presume

[14]The VAR assumption is often adopted in past literature studying return predictability (e.g., Campbell and Shiller (1988), Campbell and Vuolteenaho (2004), and Campbell et al. (2013)).

that the time-varying risk prices of the shocks to latent factor are affine in the state variables $\boldsymbol{x}_t$. However, our paper is distinct from theirs in the following aspects. First and foremost, we aim to estimate the term structure of risk premia for all (traded and nontraded) factors of equilibrium models, whereas Kelly et al. (2023) focus on dividend yields. Second, Kelly et al. (2023) specify the dynamics of asset prices and reverse-engineer the dynamics for dividend growth using the restrictions implied by the former. Conversely, our paper specifies a MA representation for dividend growth that always exists, as in equation (20). Our modelling choice is analogous to most macrofinance models that directly specify the dynamics of, e.g., consumption and dividend growth, but we do so in a general and flexible way via the MA representation.

3 Simulations

This section studies the finite-sample properties of our Bayesian estimator in Proposition 1 via Monte Carlo simulations. Throughout the simulations, we consider two sample sizes, $T \in \{200, 600\}$, matching the quarterly and monthly frequencies, respectively. We simulate asset returns from a five-factor model as in equations (1) and (2), as follows:

$$\boldsymbol{r}_t = \hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\beta}}_{\tilde{v}}\hat{\boldsymbol{\lambda}}_{\tilde{v}} - \frac{1}{2}\hat{\boldsymbol{\Upsilon}}_r + \hat{\boldsymbol{\beta}}_{\tilde{v}}\tilde{\boldsymbol{v}}_t + \boldsymbol{w}_{rt}, \quad \tilde{\boldsymbol{v}}_t \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}_K, \boldsymbol{I}_K).$$

Specifically, $\boldsymbol{r}_t$ contain Fama-French 275 portfolio returns (FF275, see Internet Appendix IA.3), and factor loadings $\hat{\boldsymbol{\beta}}_{\tilde{v}}$ are calibrated as the eigenvectors corresponding to the five largest eigenvalues of the sample covariance matrix of $\boldsymbol{r}_t$. Risk premia $\hat{\boldsymbol{\lambda}}_{\tilde{v}}$ are estimated using the observed data. To ensure that $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}_{\tilde{v}}$ are orthogonal in simulations, we regress the estimated $\boldsymbol{\alpha}$ on $\boldsymbol{\beta}_{\tilde{v}}$ and extract the residual term, denoted by $\hat{\boldsymbol{\alpha}}$. We allow for a non-diagonal covariance matrix of $\boldsymbol{w}_{rt}$. Following Bai and Ng (2002), we simulate w_{irt} as follows:

$$w_{irt} = \hat{\sigma}_{irt} \cdot \Big[e_{it} + \sum_{j\neq 0, j=-J}^{J} \beta e_{i-j,t}\Big], \quad e_{it} \overset{\text{iid}}{\sim} \mathcal{N}(0, \frac{1}{1+2J\beta^2}), \tag{23}$$

where $J = \max\{10, \text{int}\big(N/20\big)\}$, $\beta = 0.1$,[15] and $\{\hat{\sigma}^2_{ir}\}_{i=1}^{N}$ are the estimated variance of idiosyncratic shocks for each asset.

Second, we simulate strong factors. For $T = 200$, we use nondurable consumption growth

[15] β cannot be too large since we need to ensure that the largest eigenvalue of $\hat{\boldsymbol{\Sigma}}_{wr}$ is less than the smallest eigenvalue of $\hat{\boldsymbol{\beta}}_{\tilde{v}}^\top \hat{\boldsymbol{\beta}}_{\tilde{v}}$. Otherwise, some common factors cannot be identified.

to estimate impulse responses, denoted by $\{\hat{\rho}_s\}_{s=0}^{\bar{S}}$, assuming the true $\bar{S} = 8$ (quarters). For $T = 600$, we use monthly industrial production growth to obtain $\{\hat{\rho}_s\}_{s=0}^{\bar{S}}$, and the true $\bar{S}$ is 16 (months). With these parameters, we simulate the strong g_t as follows:

$$g_t = c \cdot \sum_{s=0}^{\bar{S}} \hat{\rho}_s f_{t-s} + w_{gt}, \quad f_t = \frac{1}{\sqrt{3}}(\tilde{v}_{1t} + \tilde{v}_{3t} + \tilde{v}_{5t}), \quad w_{gt} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2_{wg}), \tag{24}$$

where f_t relates to both large and small principal components (PCs) of asset returns. We consider different signal-to-noise ratios summarized by the time series fit $R^2_g \in \{30\%, 20\%, 10\%\}$.

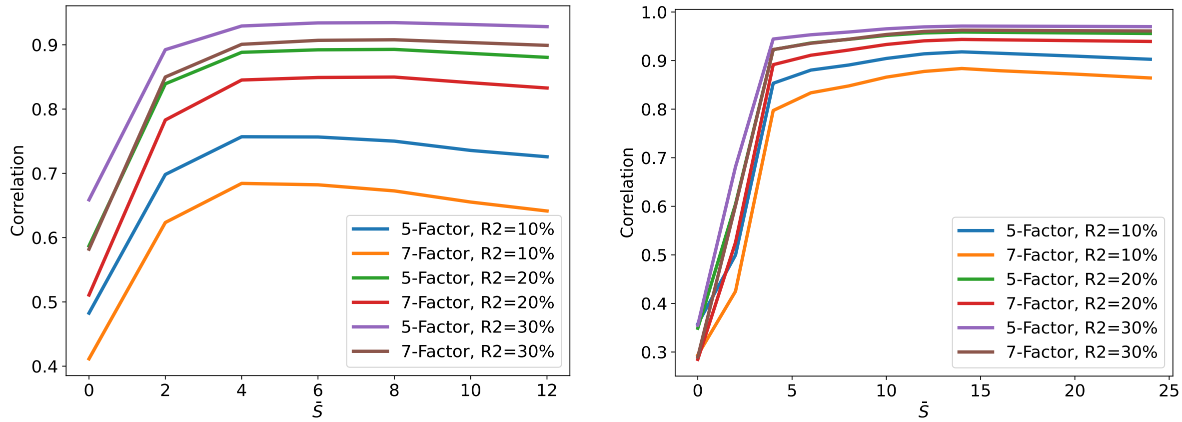
Finally, for the weak factor, we simulate f_t independently from the standard normal distribution. Nevertheless, the simulated weak factor g_t is autocorrelated, so we can use it to explore whether the Newey and West (1987) type of sandwich covariance matrix can deliver proper Bayesian credible intervals for factors with an autocorrelated measurement error.

Tables A1 and IA.III of the Internet Appendix report the empirical size of our test for strong factors in 1,000 simulations. We estimate the term structure of g_t's risk premia using $\bar{S} = 12$ for $T = 200$ and $\bar{S} = 24$ for $T = 600$. Our method provides appropriate credible intervals for g_t's risk premia as long as we include all priced latent factors in the estimation ($K \geq 5$), even in an environment with a low signal-to-noise ratio and a small sample size. However, if we omit some priced factors (e.g., the number of factors is four), our Bayesian estimates are biased because the simulated g_t loads on the fifth PC of asset returns. Nevertheless, including more factors than in the pseudo-true model has no sizable detrimental effect, suggesting that such an approach is conservative.

Can we recover the priced information embedded in g_t if we consider only the contemporaneous correlation between g_t and asset returns? To answer this question, we estimate the models with different numbers of lags $\bar{S}$. Figure 1 plots the average correlation between the true f_t and its estimate, $\hat{f}_t = \hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{v}}_t$. When we project g_t only on contemporaneous asset return shocks ($\bar{S} = 0$ in equation (8)), $\text{corr}(f_t, \hat{f}_t)$ is small, ranging from 0.4 to 0.65. As we include more lagged asset pricing information in g_t, this correlation coefficient significantly increases; hence, including the lagged asset return information is essential in identifying the priced shock driving the nontradable factor. Notably, the detrimental effect of including more lags than in the pseudo-true specification is generally very small.

Figure 2 reports the power of rejecting zero risk premia of strong factors.[16] The model with

[16]We report the power for $R^2_g \in \{10\%, 20\%\}$ in Figure IA.1 of the Internet Appendix.



(a) $T = 200$ **(b)** $T = 600$

Figure 1: Posterior median of correlation coefficients between true and estimated f_t

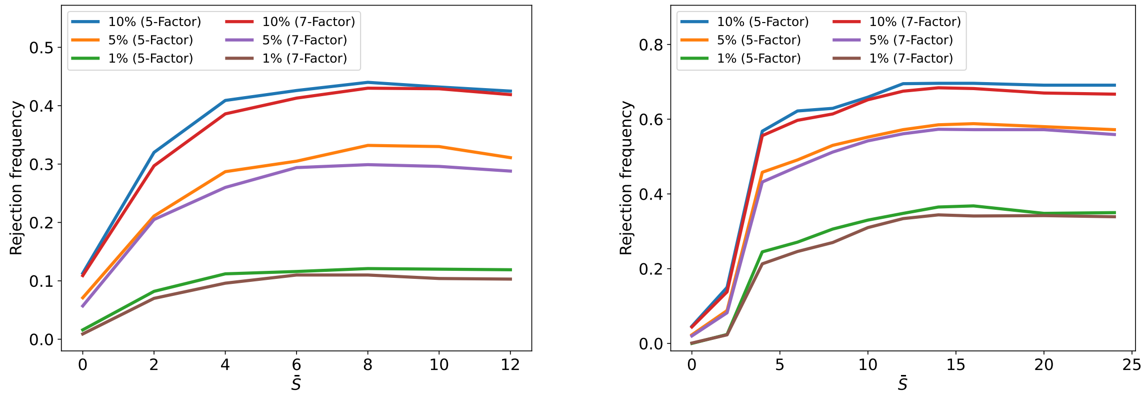
The figure plots the average $\text{corr}(\hat{f}_t, f_t)$ in 1,000 simulations, where $\text{corr}(\hat{f}_t, f_t)$ quantifies the correlation between the true f_t and its estimate, $\hat{f}_t = \hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{v}}_t$. We consider strong factors, with $R_g^2 \in \{10\%, 20\%, 30\%\}$, and two sample sizes, $T \in \{200, 600\}$. In each simulated scenario, we estimate several model configurations with different numbers of factors and different $\bar{S}$.

$\bar{S} = 0$ generally has low test power. In contrast, as we include more lagged latent factors in g_t's estimation, we considerably increase the test power. Hence, our proposed MA representation of g_t is the key to detecting significant risk premia in persistent factors.

Including more factors (e.g., in the seven-factor models) tends to be a conservative strategy since it delivers proper yet wider credible intervals of risk premia estimates. Nevertheless, it comes at the cost of lowering the power of the test and the correlation between the true and estimated f_t. In the empirical application, we will explore whether our risk premia estimates are robust to adding more latent factors, acknowledging that more factors will increase the estimation uncertainty mechanically.

In Tables IA.IV and IA.V of the Internet Appendix, we investigate useless factors that do not correlate with asset returns. The useless factors are assumed to be persistent, and a larger R_g^2 corresponds to a more persistent process. Past literature (e.g., Kan and Zhang (1999a,b)) points out the fragility of Fama-MacBeth and GMM estimates of risk premia in the presence of useless factors. It is worth noting that our Bayesian estimates do not suffer from this issue. The Bayesian credible intervals of useless factors' risk premia tend to be conservative, leading to a slight under-rejections of zero risk premia in small sample.

One potential concern is that including many lags of multiple latent factors might lead to severe overfitting of the data. To alleviate this concern, we report in the Internet Appendix (see



(a) $R_g^2 = 30\%$ & $T = 200$ **(b)** $R_g^2 = 30\%$ & $T = 600$

Figure 2: Power of identifying strong factors

The figure plots the frequency of rejecting the null hypothesis $H_0 : \lambda_g^{\bar{S}} = 0$, based on the 90%, 95%, and 99% credible intervals based on our Bayesian estimates in Proposition 1. $\lambda_g^{\bar{S}}$ is defined in equation (6). We consider strong factors, with $R_g^2 = 30\%$, and two sample sizes, $T \in \{200, 600\}$. In each simulated scenario, we estimate several model configurations with different numbers of factors and different $\bar{S}$. The number of Monte Carlo simulations is 1,000.

Table IA.VI) the posterior means of R_g^2 in 1,000 simulations. Our simulation results suggest that the posterior means of R_g^2 are reasonably close to their pseudo-true values. Hence, our approach does not lead to significantly inflated time series fits for g_t.

Moreover, we explore the performance of our Bayesian estimates for factors that correlate with only the contemporaneous asset return shocks (i.e., $\bar{S} = 0$ in the true data-generating process of g_t), which fits the model configuration studied in Giglio and Xiu (2021). Table IA.VII of the Internet Appendix shows that our Bayesian estimator has almost identical size and power to the frequentist test in Giglio and Xiu (2021) in the special case of $\bar{S} = 0$.

Finally, in Internet Appendix IA.2, we repeat our simulation study to examine the time-varying risk premia and their term structures as described in Section 2.2. Overall, size and power, as well as the correlation between filtered and calibrated latent processes (see Tables IA.VIII–IA.X in the Appendix), are similar to those reported in this section. Despite the significant added generality by modeling the latent systematic risk drivers as following a VAR(1) process, we observe only a minimal degree of attenuation bias and increased posterior uncertainty for the estimated term structure of risk premia.

4 Empirical Analysis

In this section, we apply our Bayesian framework to investigate whether factors are priced, the term structure of the factors' risk premia, and their connection to forward equity yields.

4.1 Unconditional Risk Premia in Equity Markets

We begin our empirical investigation with the unconditional risk premia in equity markets. Our analysis relies on a large cross-section of FF275, covering the period between Q3 1963 and Q4 2019. Throughout our paper, we standardize the tested factors to have unit variances per period. Definition, sample periods, and data sources of factors and test assets can be found in Internet Appendix IA.3.

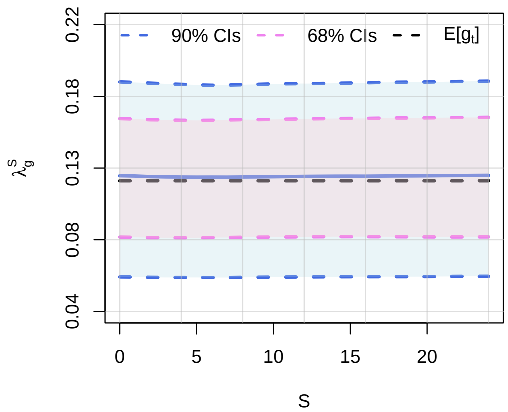
To conduct our Bayesian estimation in Section 2, we need to determine the number of latent factors, K. We adopt the selection approach proposed by Giglio and Xiu (2021)[17] and estimate that the number of factors is five in FF275 at monthly or quarterly frequencies.

Moreover, we find that the first several latent factors explain most of the time series and cross-sectional variations. In the time series dimension, the first five PCs account for more than 93% of time series variations at monthly and quarterly frequencies. Adding the 6th and 7th PCs only marginally improves the time series fit. In the cross-sectional dimension, the five-, six, and seven-factor models explain 55.0%, 58.6%, and 58.7% (59.0%, 59.3%, and 72.9%) of cross-sectional variations in average returns at the quarterly (monthly) frequency. Therefore, the statistical test in Giglio and Xiu (2021), as well as time series and cross-sectional fit, indicate that the five-factor model is a reasonable benchmark; we thus adopt it in our baseline estimations (but also conduct robustness checks with $K = 6$ or 7).

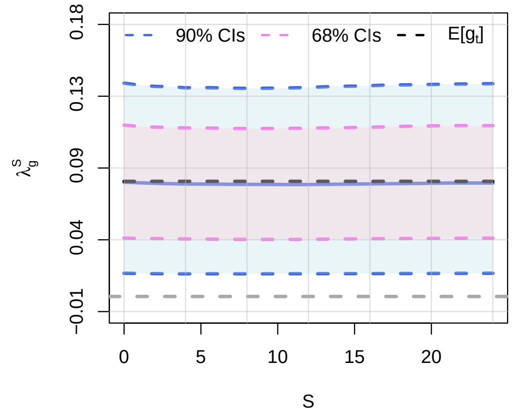
4.1.1 Term Structure of Risk Premia

We first explore Bayesian risk premia estimates of some canonical tradable factors and compare them with their time series average excess returns. Figure 3 plots the term structure of risk premia for Carhart (1997) four factors, whose risk premia are estimated using Proposition 1

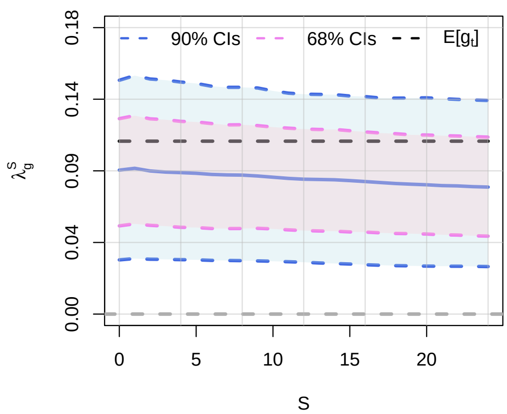
[17] We follow the method in Internet Appendix I.1 of Giglio and Xiu (2021). That is, the selected number of factors is equal to $\hat{K} = \arg\min_{1 \leq j \leq K_{\max}} \left[N^{-1}T^{-1}\gamma_j(\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{R}}) + j \times \phi(n,T)\right] - 1$, where $\bar{\boldsymbol{R}}$ is a $T \times N$ matrix of demeaned asset returns, $\gamma_j(\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{R}})$ is the j-th eigenvalue of $\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{R}}$, $\phi(n,T) = 0.5 \times \hat{\gamma} \times (\log(N) + \log(T))(N^{-\frac{1}{2}} + T^{-\frac{1}{2}})$, and $\hat{\gamma}$ is the median of the first $K_{\max}$ eigenvalues of $\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{R}}$. We set $K_{\max}$ to 15.



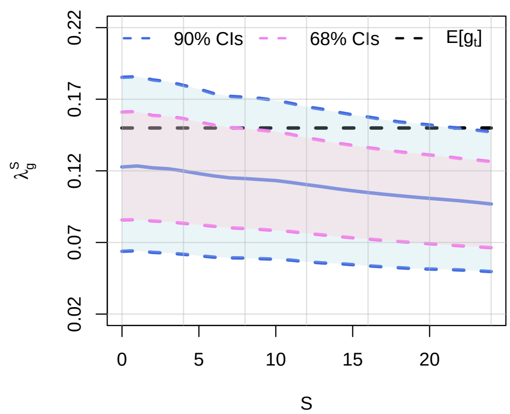
(a) MKT



(b) SMB



(c) HML



(d) MOM

Figure 3: Term structure of risk premia: Carhart four factors

Term structure of risk premia estimates (in Sharpe ratio units) using Proposition 1. The risk premium at horizon S (λ_g^S) is defined in equation (6). The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-factor models for asset returns. We study monthly Carhart (1997) four factors, whose risk premia are estimated using a lag of 24 months in g_t's equations. We include their in-sample monthly Sharpe ratios (grey dotted lines). In addition to the point estimates, we report the 68% and 90% Bayesian credible intervals, highlighted in pink and blue, respectively. Definition and data sources of factors and test assets can be found in Appendix IA.3. Sample: July 1963 to December 2019.

($\bar{S} = 24$ and $K = 5$). These tradable factors tend to have almost flat term structures of risk premia. The Bayesian point estimates (solid blue lines) have similar magnitudes as the time series Sharpe ratios (grey dotted lines), which are covered by the 68% Bayesian credible intervals (purple dotted lines). Therefore, our approach provides estimates very close to the time series averages of tradable factors in both economic and statistical sense.

Next, we study other economic variables and report their term structures of risk premia estimates in Table 1. For quarterly (monthly) variables, we conduct the Bayesian estimation

Table 1: Factors' risk premia: Five-factor models

Panel A. Quarterly variables, $\bar{S} = 12$ quarters								
$S =$	0	2	4	6	8	10	12	R_g^2
AEM intermediary	0.082***	0.077**	0.078**	0.063	0.046	0.026	0.019	14.6%
Capital share growth	0.008	0.009	0.005	0.001	-0.003	-0.007	-0.012	7.4%
GDP growth	0.026*	0.084***	0.133***	0.164***	0.180***	0.195***	0.204***	23.4%
IP growth	0.008	0.087***	0.145***	0.177***	0.194***	0.204***	0.205***	37.5%
Durable consumption growth	-0.014	0.079**	0.122***	0.140***	0.147***	0.153***	0.158***	18.3%
Nondurable consumption growth	0.042***	0.103***	0.141***	0.179***	0.206***	0.226***	0.244***	22.4%
Service consumption growth	0.006	0.013	0.020	0.027	0.035	0.041	0.045	9.8%
Nondurable + service	0.028*	0.067*	0.099*	0.127**	0.148*	0.163**	0.181**	18.5%
Labor income growth	0.000	0.002	0.003	0.004	0.005	0.005	0.009	5.3%
Dividend growth	0.009	0.045*	0.109***	0.175***	0.245***	0.306***	0.357***	40.3%
Macro PC1 (FRED-QD)	0.019	0.092***	0.165***	0.222***	0.266***	0.303***	0.332***	47.9%
Macro PC2 (FRED-QD)	0.098***	0.147***	0.148**	0.129*	0.110	0.091	0.070	37.1%
Macro PC3 (FRED-QD)	-0.003	-0.003	-0.002	-0.002	-0.001	0.000	0.001	10.8%
Macro PC4 (FRED-QD)	-0.151***	-0.173***	-0.226***	-0.286***	-0.341***	-0.389***	-0.434***	47.3%
Macro PC5 (FRED-QD)	0.050	0.055	0.044	0.029	0.018	0.011	0.003	29.3%
Panel B. Monthly variables, $\bar{S} = 24$ months								
$S =$	0	4	8	12	16	20	24	R_g^2
Oil price change	-0.004	-0.023	-0.034	-0.039	-0.042	-0.041	-0.040	7.1%
TED spread change	0.000	-0.001	0.000	0.000	0.001	0.001	0.001	8.8%
Nontraded HKM intermediary	0.098***	0.101***	0.097***	0.093***	0.091***	0.089***	0.088***	60.8%
Traded HKM intermediary	0.114***	0.115***	0.110***	0.104***	0.100***	0.098***	0.096***	71.0%
PS liquidity	0.050***	0.074***	0.086***	0.097***	0.108***	0.118***	0.126***	15.0%
$\Delta\log(\text{VIX})$	-0.131***	-0.079***	-0.062***	-0.049***	-0.042***	-0.037***	-0.032***	51.6%

The table reports Bayesian estimates of factors' risk premia using Proposition 1, where the risk premia over S horizons (λ_g^S) are defined in equation (6). The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider a five-factor model for asset returns. Panel A tabulates the estimates of quarterly factors, using a lag of 12 quarters in g_t's equations. Panel B tabulates the estimates of monthly factors, using a lag of 24 months in estimation. We use Bayesian credible intervals to conduct hypothesis testing: If the 90% (95%, 99%) credible interval of g_t's risk premium does not contain zero, the risk premium estimate will be highlighted by * (**, ***). Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

as in Proposition 1, using a lag of 12 quarters (24 months) in g_t's equations. Four empirical findings in Table 1 are noteworthy.

First, many macro factors carry significant risk premia, including IP growth, GDP growth, durable and nondurable consumption growth, dividend growth, and macro PCs 1, 2, and 4 in the FRED-QD dataset of McCracken and Ng (2020).[18] More interestingly, most of them have *upward-sloping* term structures of risk premia, as shown in Figure 4. At quarterly frequency ($S = 0$), most macroeconomic factors are weakly identified at best. However, risk premia carried by these macro factors are significant and as large as that of the market at business cycle frequencies (two to three years). Therefore, these macro factors are riskier from the perspective

[18] Dividend growth is the quarterly growth of the smoothed aggregate dividend payments made in the previous 12 months. We consider the smoothed annual dividends of the S&P 500 index in order to remove the mechanical seasonality in the dividend payments.

of long-term than short-term investors. The only exception among the priced macro factors is macro PC2, where we detect an almost flat term structure.

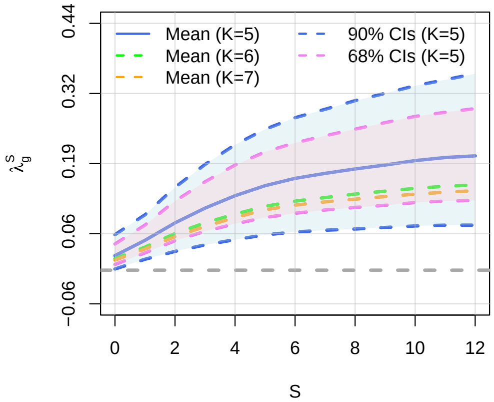
Second, the observations in Table 1 have direct implications for leading macro-finance models. Figure A1 of the Appendix plots the term structure of risk premia in the habit (Campbell and Cochrane (1999)) and long-run risk frameworks (Bansal and Yaron (2004)).[19] Specifically, the habit model implies a flat term structure of consumption risk premia, whereas it is upward-sloping in the long-run risk model. With respect to dividend growth, we consider the quarterly growth of the smoothed dividend payment (defined as the aggregate dividend payments made in the previous 12 months) to be consistent with our empirical analysis. Even though both models predict upward-sloping term structures of risk premia for smoothed dividend growth, the magnitudes and slopes are much more sizable in the long-run risk model than in the habit model. Overall, the long-run risk model tends to be more consistent with our estimates for nondurable consumption and dividend growth.

Third, our findings are not a simple byproduct of factor persistence. For instance, durable consumption growth, Adrian et al. (2014) (AEM) intermediary factor, and labour income growth have similar autocorrelation structures. However, as we show in Table 1, their term structures of risk premia are totally different: upward-sloping for durable consumption growth, slightly downward-sloping for AEM intermediary factor, and flat for labour income growth. Therefore, the term structure of risk premia is driven by the propagating mechanism of how the economic factor responds to the f_t shock over time (rather than just its persistence).

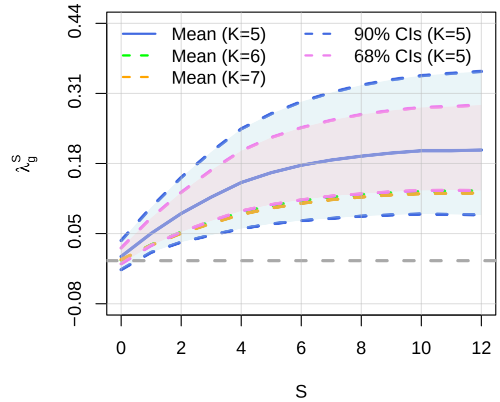
Fourth, the term structure of VIX risk premia (more precisely, their absolute values) is downward-sloping. The mimicking portfolio hedging against monthly VIX changes earns a sizable risk premium of -0.13, but the two-year risk premium declines to only -0.03, although still significant. This observation is consistent with the previous literature (Eraker and Wu (2014), Dew-Becker et al. (2017), and Johnson (2017)), which estimates VIX risk premia using derivative contracts with different expiration dates.

We further confirm that we can interpret the term structure of risk premia estimates from the angle of horizon-specific mimicking portfolios. Figure IA.2 in the Internet Appendix plots the per-period mean returns of the horizon-specific mimicking portfolios hedging against the six macro factors in Figure 4. As we show therein, these portfolios display increasing term structures of risk premia that are similar to what we find in Figure 4.

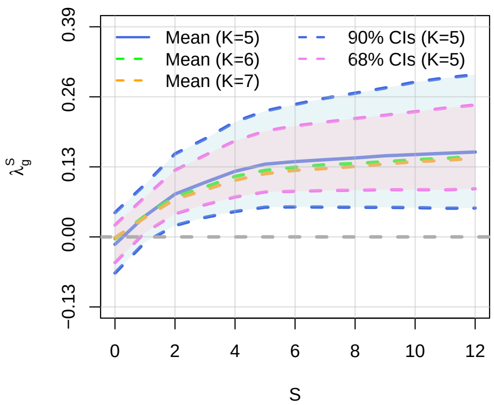
[19] We discuss the calibrations in detail in Internet Appendix IA.4.



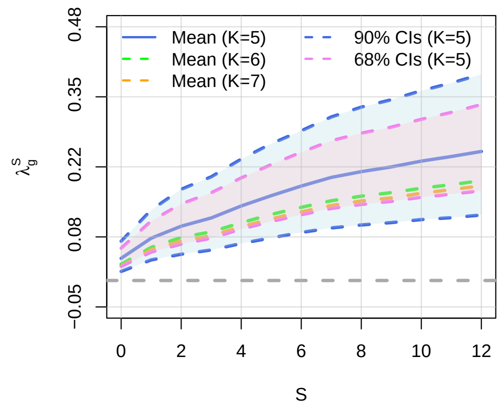
(a) GDP growth



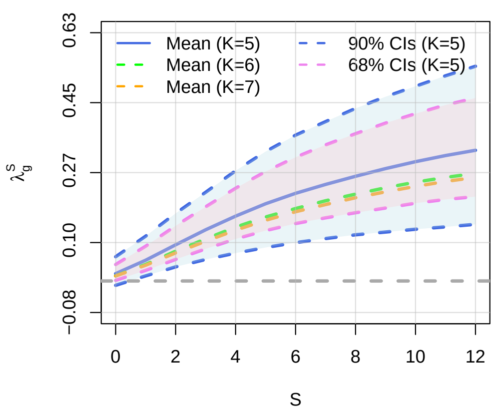
(b) IP growth



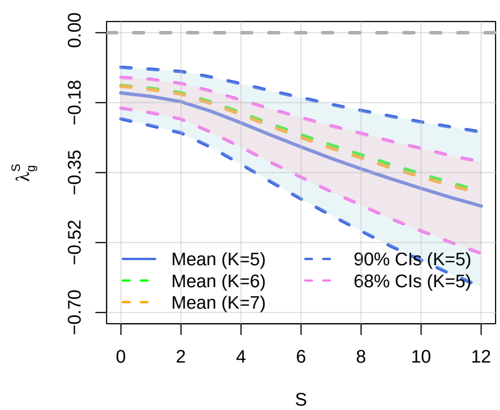
(c) Durable consumption growth



(d) Nondurable consumption growth



(e) Macro PC1 (FRED-QD)

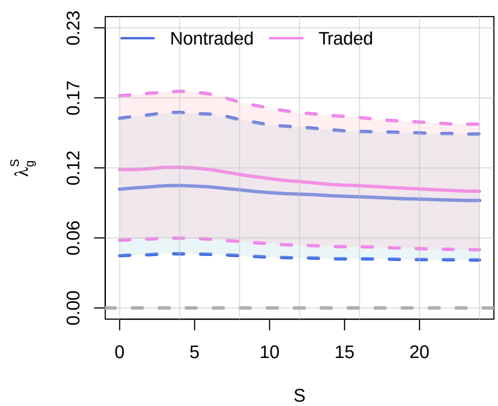


(f) Macro PC4 (FRED-QD)

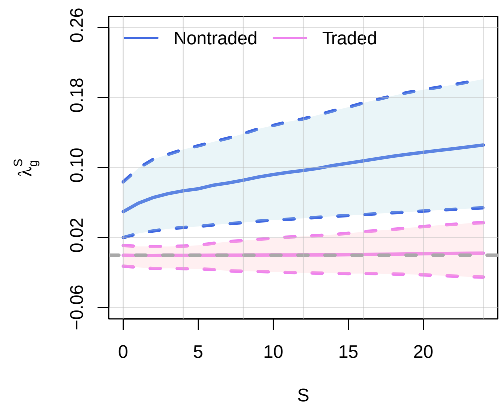
Figure 4: Term structure of factor's risk premia: Some priced macro factors

This figure plots the term structure of risk premia estimates using Proposition 1, where the risk premium over S horizons (λ_g^S) is defined in equation (6). The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-, six- and seven-factor models for asset returns. In addition to the point estimates, we show the 68% and 90% Bayesian credible intervals based on five-factor models, highlighted in pink and blue, respectively. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

However, simple mimicking portfolios based on single-period risk exposures may fail to capture the entire term structure of risk premia embedded in economic factors. Figure 5 plots the estimates for both traded and nontraded versions of the He et al. (2017) (HKM) intermediary factors (Panel (a)) and Pástor and Stambaugh (2003) (PS) liquidity factors (Panel (b)). The HKM traded and nontraded factors command almost the same risk premia across different horizons. The term structures are almost flat, so the nontraded HKM risk factor has an almost zero forward beta (see the discussion in Example 3). Conversely, the tradable version of the PS liquidity factor, which ignores the positive forward betas, fails to capture the upward-sloping term structure.



(a) HKM intermediary factors



(b) PS liquidity factors

Figure 5: Term structure of factor's risk premia: Traded vs. Nontraded versions

This figure plots the term structure of risk premia estimates for both traded and nontraded versions of the He et al. (2017) intermediary factors and Pástor and Stambaugh (2003) liquidity factors. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-factor models for asset returns. In addition to the point estimates, we show the 90% Bayesian credible intervals. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

4.1.2 Risk Price of the f_t Shock to Nontraded Factors

We go on to explore the role of the f_t shock in the latent SDF. We use the SDF representation in Remark 1; that is, $m_t = 1 - \lambda_f f_t - \boldsymbol{\lambda}_u^\top \boldsymbol{\Sigma}_u^{-1} \boldsymbol{u}_t$, where $\boldsymbol{u}_t$ are orthogonal to f_t and act as the control for omitted sources of priced risk. Table 2 reports the risk price estimates of f_t for several priced nontraded risk factors based on the evidence in Table 1. We show that these f_t shocks are indeed priced in the cross-section. Furthermore, the annualised Sharpe ratios implied by these f_t shocks are economically large yet not excessive — 0.42 to 0.71 per year — on par with that of the market index. Finally, the column $\mathbb{E}\left[SR_f^2/SR_m^2 \mid \text{data}\right]$ quantifies the importance of f_t in the latent SDF. We find that these economic sources of risk explain individually about 13 − 58% of the SDF's variance. Therefore, a significant amount of priced shocks in financial markets are not captured by these economic factors, further highlighting the importance of controlling for omitted variables in the estimation.

Table 2: Risk price of the f_t shock to nontraded risk factors

	λ_f	$\mathbb{E}[SR_f \mid \text{data}]$	$\mathbb{E}\left[\frac{SR_f^2}{SR_m^2} \mid \text{data}\right]$
Panel A. Quarterly variables, $\bar{S} = 12$ quarters			
AEM intermediary	0.356 [0.205, 0.494]	0.711 [0.412, 0.987]	0.580 [0.225, 0.858]
GDP growth	0.226 [0.098, 0.353]	0.452 [0.196, 0.706]	0.231 [0.047, 0.513]
IP growth	0.223 [0.096, 0.341]	0.447 [0.192, 0.682]	0.221 [0.046, 0.479]
Durable consumption growth	0.340 [0.190, 0.472]	0.681 [0.393, 0.944]	0.528 [0.201, 0.826]
Nondurable consumption growth	0.283 [0.157, 0.411]	0.567 [0.314, 0.822]	0.358 [0.122, 0.660]
Nondurable + service	0.212 [0.046, 0.369]	0.425 [0.114, 0.740]	0.204 [0.015, 0.558]
Dividend growth	0.245 [0.120, 0.375]	0.491 [0.240, 0.750]	0.270 [0.072, 0.558]
Macro PC1 (FRED-QD)	0.213 [0.098, 0.329]	0.427 [0.196, 0.658]	0.204 [0.047, 0.444]
Macro PC4 (FRED-QD)	-0.294 [-0.413, -0.173]	0.589 [0.346, 0.827]	0.385 [0.158, 0.637]
Panel B. Monthly variables, $\bar{S} = 24$ months			
Nontraded HKM intermediary	0.126 [0.056, 0.201]	0.437 [0.193, 0.696]	0.136 [0.028, 0.314]
PS liquidity	0.152 [0.067, 0.226]	0.528 [0.231, 0.783]	0.196 [0.040, 0.405]
$\Delta \log(\text{VIX})$	-0.148 [-0.259, 0.251]	0.655 [0.371, 0.950]	0.327 [0.115, 0.595]

The table reports (1) the risk price of the f_t shock to the nontraded risk factors (column λ_f), (2) the annualized Sharpe ratio implied by the $\lambda_f f_t$ component (column $\mathbb{E}[SR_f \mid \text{data}]$), and (3) the share of SDF variance explained by f_t (column $\mathbb{E}\left[SR_f^2/SR_m^2 \mid \text{data}\right]$), based on the same estimates as in Table 1 and the SDF representation in Remark 1. In each column, we report both the posterior median and the 90% posterior credible intervals.

4.1.3 Contemporaneous Innovations in Macro Factors

Empirically, researchers often fail to identify priced macro risks when studying only the contemporary correlations between asset returns and macro factors. The first column of Table 1, and Figure 4, indicate that the risk premia of GDP growth, IP growth, durable consumption growth, dividend growth, and macro PC1 are tiny and insignificant at $S = 0$. One concern of the analysis in Table 1 and Figure 4 is that we include many lags in the estimation, leading to noisier risk premia estimates. To alleviate this concern we repeat the estimation using $\bar{S} = 0$. Panel A of Table 3 shows that among the eight priced macro factors mentioned above, only macro PC2 and PC4 carry significant risk premia in this case.

Panel B further extracts the AR(1) innovations in macro factors and estimates their risk premia by setting $\bar{S} = 0$. Similar to Panel A, we observe only macro PC2, PC4, and nondurable plus service consumption (albeit neither nondurable nor service consumption is priced at $\bar{S} = 0$)

being priced, while all other macro factors have negligible and insignificant risk premia. Although the AR(1) model is often used in both empirical and theoretical works, extracting the AR(1) innovations is insufficient to recover the risk premia of many macro variables, either because the AR(1) shocks are inconsequential or the AR(1) assumption is questionable. Differently, the MA representation does not take a stance on their exact data-generating processes. We model the priced component of the macro factors as a flexible MA of both the current and lagged asset return innovations.

But why does including lagged asset return shocks in g_t's equation enable us to identify the priced risk? The time series fit, R_g^2, sheds light on this issue. For most traditional macro factors, R_g^2 values in Table 1 are considerably larger than those in Table 3. For instance, the contemporaneous asset return shocks explain only 3% of time series variations in macro PC1, but its R_g^2 increases to 48% in the estimation with $\bar{S} = 12$ quarters, hence greatly enhancing the signal-to-noise ratio and our ability to identify the risk premia. In contrast, comparing the R_g^2 of AEM and HKM factors in Table 1 with those in Table 3, we find that lagged asset return innovations are not essential in driving intermediary factors. For these factors, estimating their risk premia using $\bar{S} = 0$ seems to be a better choice.

Remark 2. *There is extensive literature on developing new estimators of risk premia that are robust to weak factors, including Kan et al. (2013), Gospodinov et al. (2014, 2019), Bryzgalova (2015), Kleibergen and Zhan (2020), Anatolyev and Mikusheva (2022), and Bryzgalova et al. (2023). Our Bayesian estimator is not only robust to the weak identification issue but, more importantly, transforms some weak macro factors at short horizons into strongly identified ones at business-cycle frequencies. With this regard, we successfully recover the priced risk in macro variables through the lens of horizon-specific risk.*

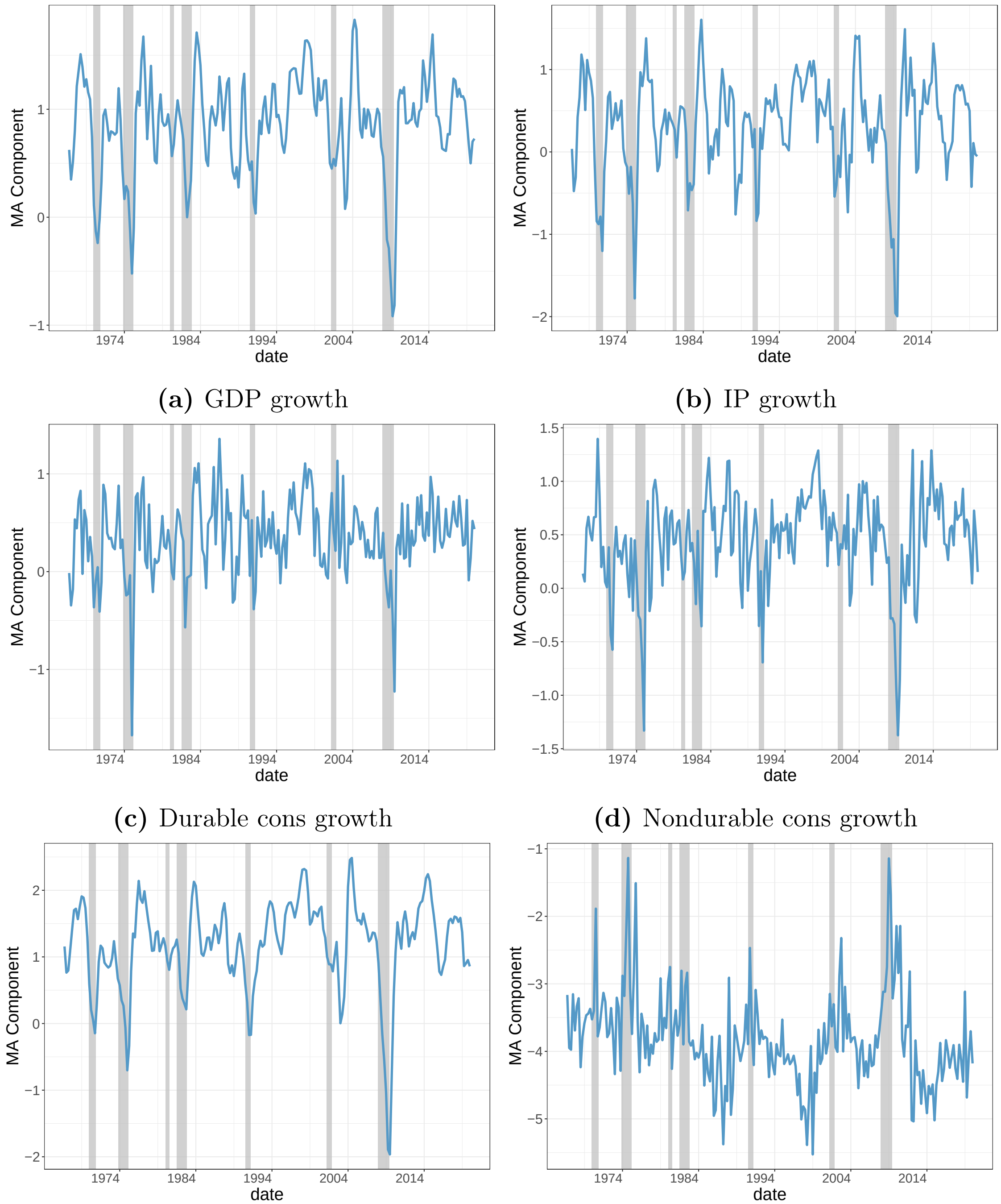
4.1.4 MA Components of Macro Factors

Perhaps the most surprising empirical finding is that macro variables carry much more sizeable risk premia at long horizons ($S = 8$ to 12 quarters) than at quarterly frequency ($S = 0$). What is the economic intuition behind this phenomenon? To help answer this question, we plot in Figure 6 the MA component spanned by six priced macro variables and asset return factors, that is, $\sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top \boldsymbol{v}_{t-s}$. Strikingly, the MA components of all these six macro variables present clear business cycle patterns — long-horizon investors who hedge against low (high) realizations of these macro factors require positive (negative) risk premia.

Table 3: Factors' risk premia: $\bar{S} = 0$

	$\mathbb{E}[\lambda_g \mid \mathcal{D}]$			$\mathbb{E}[R_g^2 \mid \mathcal{D}]$		
Number of factors:	5	6	7	5	6	7
Panel A. Original factors						
AEM intermediary	0.141***	0.176***	0.175***	10.4%	12.2%	12.4%
Capital share growth	0.032	0.013	0.013	1.8%	2.8%	2.8%
GDP growth	0.005	0.013	0.013	4.2%	4.3%	4.3%
IP growth	-0.028	0.004	0.003	2.9%	4.3%	4.3%
Durable consumption growth	-0.011	0.000	-0.002	7.7%	7.9%	8.2%
Nondurable consumption growth	0.042	0.058	0.056	3.7%	4.1%	4.1%
Service consumption growth	0.015	0.053	0.052	4.0%	6.4%	6.5%
Nondurable + service	0.032	0.067*	0.066*	4.0%	5.9%	6.0%
Labor income growth	-0.006	0.035	0.028	1.5%	3.8%	8.7%
Dividend growth of SP500	0.037	0.037	0.044	5.1%	5.4%	11.7%
Macro PC1 (FRED-QD)	-0.009	0.019	0.021	2.8%	3.9%	4.6%
Macro PC2 (FRED-QD)	0.140***	0.109**	0.103**	21.3%	22.0%	23.0%
Macro PC3 (FRED-QD)	-0.063*	-0.076**	-0.078**	4.3%	4.7%	4.8%
Macro PC4 (FRED-QD)	-0.155***	-0.164***	-0.168***	25.2%	25.4%	26.6%
Macro PC5 (FRED-QD)	0.068	0.108**	0.101**	24.8%	25.5%	27.9%
Oil price change	-0.018	-0.017	-0.016	2.6%	4.5%	4.5%
TED spread change	-0.034	-0.040*	-0.033	6.8%	10.9%	17.4%
Nontraded HKM intermediary	0.100***	0.104***	0.104***	60.3%	61.0%	61.1%
Traded HKM intermediary	0.112***	0.116***	0.116***	70.3%	71.0%	71.2%
PS liquidity	0.061***	0.059***	0.062***	11.9%	12.2%	12.9%
$\Delta \log(\text{VIX})$	-0.120***	-0.118***	-0.118***	42.8%	43.0%	43.1%
Panel B. AR(1) shocks of macro factors						
GDP growth	0.005	0.012	0.012	4.2%	4.3%	4.3%
IP growth	-0.021	0.001	-0.002	3.6%	4.3%	4.9%
Durable consumption growth	-0.010	0.002	0.000	7.4%	7.6%	7.8%
Nondurable consumption growth	0.042	0.058	0.057	3.7%	4.1%	4.1%
Service consumption growth	0.016	0.055	0.055	3.7%	6.3%	6.4%
Nondurable + service	0.031	0.069*	0.069*	3.7%	6.0%	6.0%
Labor income growth	-0.007	0.033	0.027	1.4%	3.6%	8.3%
Dividend growth of SP500	0.067*	0.068*	0.074*	3.3%	3.4%	9.5%
Macro PC1 (FRED-QD)	0.014	0.016	0.013	6.1%	6.1%	7.1%
Macro PC2 (FRED-QD)	0.109***	0.082*	0.077	23.9%	24.8%	27.6%
Macro PC3 (FRED-QD)	-0.052	-0.051	-0.055	3.3%	3.3%	4.1%
Macro PC4 (FRED-QD)	-0.150***	-0.160***	-0.165***	26.8%	26.9%	28.7%
Macro PC5 (FRED-QD)	0.057	0.074	0.068	33.8%	33.2%	37.1%
Oil price change	-0.026	-0.025	-0.025	3.2%	4.8%	4.8%

The table reports Bayesian estimates of (1) factors' risk premia and (2) time series fit R_g^2. Panel A considers the original variables that are identical to those in Tables 1 and IA.XI, whereas Panel B studies the AR(1) shocks of some macro factors. We estimate model parameters using Proposition 1 by setting $\bar{S} = 0$. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-, six-, and seven-factor models for asset returns. For risk premia estimates, we use Bayesian credible intervals to conduct hypothesis testing: If the 90% (95%, 99%) credible interval of g_t's risk premium does not contain zero, the risk premium estimate will be highlighted by * (**, ***). Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.



(a) GDP growth

(b) IP growth

(c) Durable cons growth

(d) Nondurable cons growth

(e) Macro PC1 (FRED-QD)

(f) Macro PC4 (FRED-QD)

Figure 6: Moving average components of some macro factors

This figure plots the time series of (posterior means of) moving average components spanned by asset returns' latent factors: $\sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top \boldsymbol{v}_{t-s}$, with $\bar{S} = 12$ quarters. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-factor models for asset returns. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3. Sample: Q3 1963 to Q4 2019.

Are the MA components of the priced macro factors similar? Table 4 shows that macro PC1, GDP growth, and IP growth have highly correlated MA components, often with correlation coefficients of about 90%, and their correlation with nondurable consumption is 70% or more. Nevertheless, the MA components of other macro variables, although correlated, seem to contain considerably independent information. In short, we detect some string commonality in the priced component of these macro variables, but they are not all alike.[20]

Table 4: Are MA components of macro factors similar in five-factor models?

	GDP growth	IP growth	Durable	Nondurable	Service	Dividend	Macro PC1	Macro PC2	Macro PC4
GDP growth	1.00	0.90	0.69	0.70	0.60	0.39	0.90	0.43	-0.45
IP growth	0.90	1.00	0.72	0.70	0.57	0.35	0.85	0.40	-0.25
Durable	0.69	0.72	1.00	0.64	0.35	0.32	0.61	0.32	-0.19
Nondurable	0.70	0.70	0.64	1.00	0.59	0.48	0.73	0.34	-0.55
Service	0.60	0.57	0.35	0.59	1.00	0.43	0.72	0.13	-0.43
Dividend	0.39	0.35	0.32	0.48	0.43	1.00	0.64	-0.25	-0.59
Macro PC1	0.90	0.85	0.61	0.73	0.72	0.64	1.00	0.15	-0.51
Macro PC2	0.43	0.40	0.32	0.34	0.13	-0.25	0.15	1.00	-0.19
Macro PC4	-0.45	-0.25	-0.19	-0.55	-0.43	-0.59	-0.51	-0.19	1.00

The table reports the correlation among the moving average components spanned by asset returns' latent factors, $\sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top \boldsymbol{v}_{t-s}$, with $\bar{S} = 12$ quarters. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-factor models for asset returns. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

4.1.5 Robustness Checks: More Latent Factors

Which principal components of asset returns drive g_t? Table IA.XII in the Internet Appendix reports the posterior means of the squared correlation[21] between the common component estimates, $\hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{v}}_t$, and the first seven PCs of asset returns, where the posterior distributions of $\hat{\boldsymbol{\eta}}_g$ and $\hat{\boldsymbol{v}}_t$ are estimated using a seven-factor model. The first PC of asset returns is the most important, particularly for the priced factors. Specifically, PC1 of asset returns accounts for 56–87% of the time series variations in the common components of GDP growth, IP growth, nondurable consumption growth, dividend growth, macro PCs 1, 2, 4, HKM intermediary factors, the liquidity factor, and the VIX changes. Overall, the common component is spanned mainly by the first five PCs of asset returns.

However, several variables are closely related to PC6 and PC7 of equity portfolio returns. For example, these two small PCs explain 48% of the common component in labor income growth. Furthermore, PC6 of asset returns accounts for 28%, 28%, and 9% of common components in

[20] Table IA.XIII in the Internet Appendix repeats these analyses in six- and seven-factor models, showing very similar empirical patterns.

[21] We do not report the correlation since we cannot identify the sign of $\hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{v}}_t$.

capital share growth, macro PC3, and oil price change. While labor income growth, capital share growth, and macro PC3 are not priced in six- and seven-factor models, the risk premia estimates of oil price change become significantly negative after we include PC6 of asset returns. Therefore, it is important to conduct robustness checks by considering different numbers of latent factors. We report the term structure of risk premia estimates based on six- and seven-factor models in the Internet Appendix. (See Table IA.XI) The point estimates of most factors are nearly unchanged, but Bayesian credible intervals often become wider, consistent with the observations in simulation studies.

4.2 Time-Varying Term Structure of Macroeconomic Risk Premia

We now turn to the analysis of the time variation in the term structure of macroeconomic factors' risk premia, applying the method in Section 2.2. Since the dynamics of latent factors, $\boldsymbol{v}_t$, determines the time variation in factor risk premia, we first investigate whether the five largest PCs of asset returns can be predicted by their one-period lags and other external economic variables. Following past literature (e.g., Campbell and Vuolteenaho (2004), Campbell et al. (2013), and Gagliardini et al. (2016)), we include as external predictors the price-earning ratio as well as term, default, and value spreads.

Table IA.XIV in the Internet Appendix shows that external predictors have limited predictive power. In Panel A, we consider only the four external predictors. Although value and term spread can predict PC1 and PC5 to a certain extent, the adjusted R^2s are very smalll or even negative in these specifications. We further include the lagged return PCs in Panel B and observe economically sizable predictability. For example, the adjusted R^2 is above 11% for PC4 at the quarterly frequency. In contrast, all external predictors are almost inessential in these regressions. Therefore, using them to model time-varying risk premia will introduce huge estimation noise, which can lead to attenuation bias in risk premia estimates.

Using the VAR(1) formulation for the latent systematic factors, we estimate the term structure of unconditional risk premia for the same set of variables as in Table 1. Figures IA.6–IA.8 shows the empirical results, in which the blue lines and shaded areas present the estimates based on the conditional models.[22] For comparison, we also include the previous estimates (the

[22] In particular, in Figure IA.6, the VAR(1) model contains only the latent factors of returns, whereas, in Figure IA.8, both latent factors and external predictors are included in the VAR(1) model. Unlike these two models, we impose a restriction in Figure IA.7 that both the latent factors and external predictors are driven only by the lagged external predictors.

purple lines and areas) in Table 1 based on the unconditional models. The point estimates are almost identical in both conditional and unconditional models, although we occasionally detect some minor attenuations and wider confidence intervals due to the additional parameters in the VAR system. Overall, the risk premia estimates based on the unconditional models are able to deliver consistent estimates even if the true model is time-varying.

Having established the robustness of the unconditional risk premia estimates, we proceed to explore the time-varying term structure of macro risk premia. Figure 7 reports the posterior means of the risk premia at one-quarter to three-year horizons for nondurable consumption, GDP, and industrial production growths (in Panels (a)–(c), respectively), using four external predictors to model conditional factors' risk premia. The figure highlights a clear commonality in the business cycle behavior of the term structures of macroeconomic risk premia.[23]

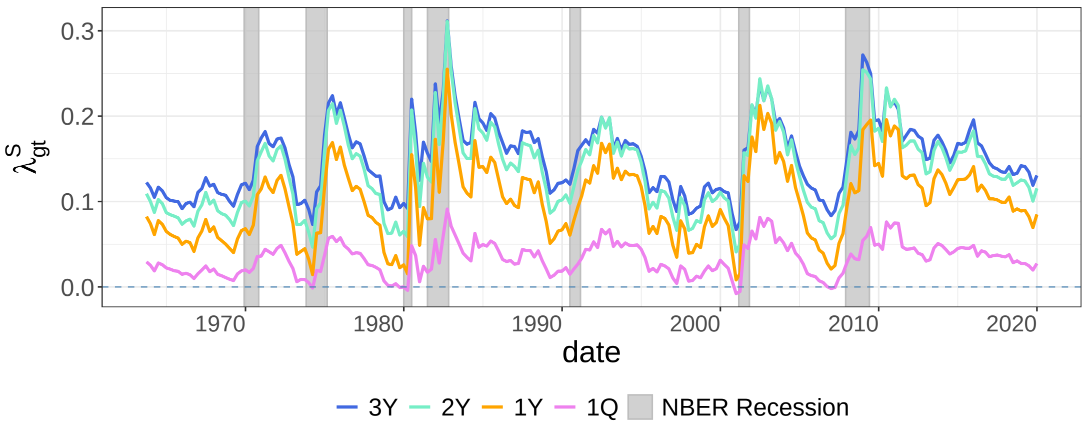
Two observations are noteworthy. First, the average level is strongly countercyclical, with smaller risk premia during expansion and a significant increase during recession episodes. Second, short-maturity (e.g., one-quarter) macro risk premia exhibit very small time variation, confirming that macroeconomic variables are weak factors at best at short horizons, even conditionally.

4.3 Term Structure of (Dividend) Risk Premia vs Strips

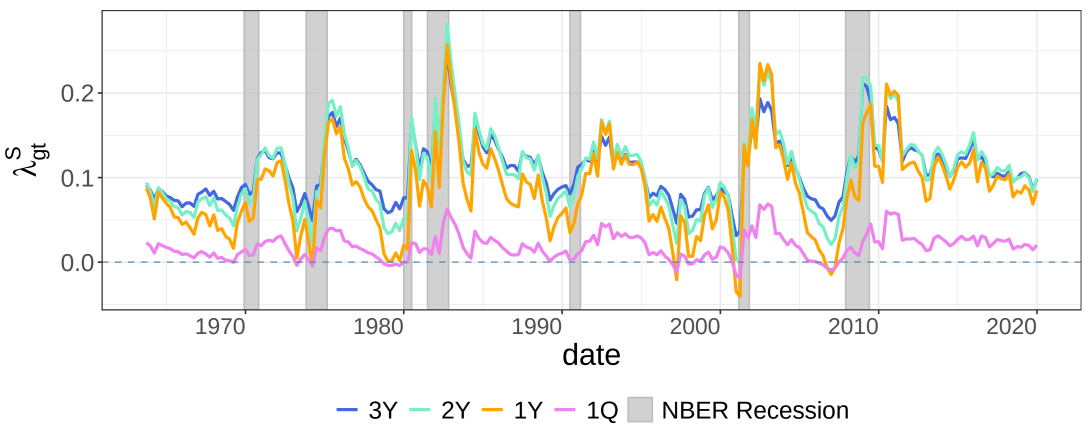
In this subsection, we study the connection between the term structure of risk premia defined in equation (21) and that of dividend strips that have been extensively studied in past literature (e.g., van Binsbergen et al. (2012), van Binsbergen and Koijen (2017), Bansal et al. (2021), and Giglio et al. (2023)). Suppose that D_t is the dividend payment at time t, and $P_{s,t}$ denotes the time-t price of the dividend strip that delivers D_{t+s} at time $t+s$. We define the holding period return on this dividend strip, as well as the spot and forward equity yield, as follows:

$$
\begin{aligned}
\text{holding period return: } R_{t,t+S} &= \frac{D_{t+s}}{P_{s,t}}, \\
\text{spot equity yield: } e_{s,t} &= \frac{1}{s}\log\left(\frac{D_t}{P_{s,t}}\right), \text{ and} \\
\text{forward equity yield: } e^f_{s,t} &= \frac{1}{s}\log\left(\frac{D_t}{P_{s,t}}\right) - \frac{1}{s} r_{f,t,t+s},
\end{aligned}
$$

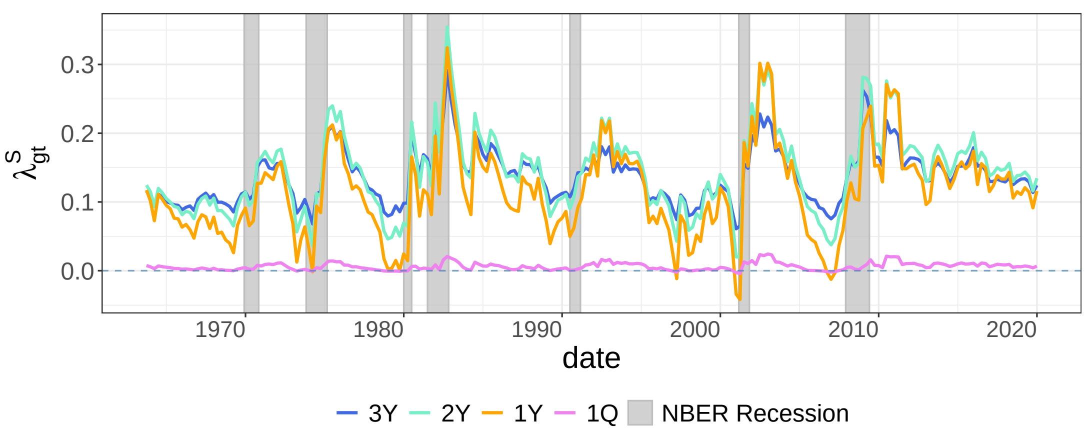
[23]Figure IA.9 in the Internet Appendix shows the time-varying risk premia for durable consumption and dividend growth, as well as the Macro PC1 and PC4. Besides, we present in Figure IA.10 the time-varying risk premia based on a different VAR(1) model that contains only the latent factors of returns.



(a) Nondurable Consumption Growth



(b) GDP Growth



(c) Industrial Production Growth

Figure 7: Time-varying term structure of macroeconomic factor's risk premia

This figure plots the time-varying term structure of risk premia following the method in Section 2.2. Risk premia of latent factors are linear in four external predictors: PE ratio of S&P 500, Term spread, default spread, and value spread. Estimates are based on the composite cross-section of 275 Fama-French characteristic-sorted portfolios. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

where $\frac{1}{s}r_{f,t,t+s}$ is log risk-free rate with time-to-maturity s.

In Internet Appendix IA.5, we show that under joint log-normality of SDF and dividend growth, forward equity yield and dividend risk premia satisfy the following relationship:

$$e^f_{s,t} = \lambda^s_{dt} - \mathbb{E}_t[g_{d,t,t+s}] - \frac{1}{2s}\text{var}_t(\Delta d_{t,t+s}) \approx \lambda^s_{dt} - \mathbb{E}_t[g_{d,t,t+s}], \tag{25}$$

where $g_{d,t,t+s} = \frac{1}{s}\log\left(\frac{D_{t+s}}{D_t}\right)$ is the per-period log dividend growth rate, $\Delta d_{t,t+s} = \log\left(\frac{D_{t+s}}{D_t}\right)$ is the multiperiod dividend growth, and $\lambda^s_{dt} = -\frac{1}{s}\text{cov}_t(m_{t,t+s}, \Delta d_{t,t+s})$ is the s-period dividend risk premium defined in equation (21). Since $\text{var}_t(\Delta d_{t,t+s})$ is empirically negligible, we can approximate the forward equity yield with $\lambda^s_{dt} - \mathbb{E}_t[g_{d,t,t+s}]$.

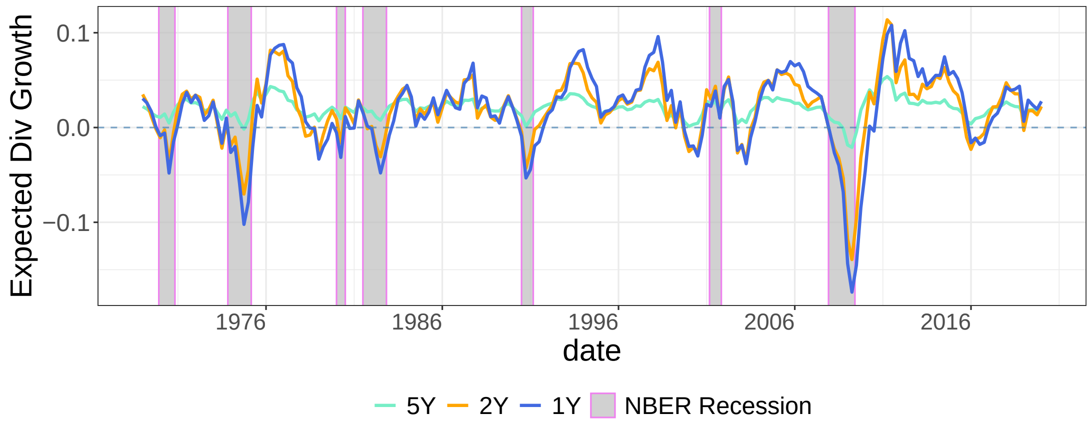
Equation (25) makes clear the distinction between the term structure of dividend risk premia and its strips: the forward equity yields are driven by both dividend risk premia and expected dividend growths. As we show in Internet Appendix IA.5, dividend risk premia can be interpreted as the per-period risk premium on the hold-to-maturity dividend strips. We estimate the term structure of unconditional dividend risk premia using our MA formulation and show in Figure IA.11 that our estimated (one- to five-year) premia, unconditionally, are extremely similar to those obtained in Bansal et al. (2021), although we use an entirely different data sample and methodology.

Nevertheless, to obtain the term structure of dividend strips, we also need to estimate the conditional mean of dividend growth. This is obtained using equation (4) as follows:

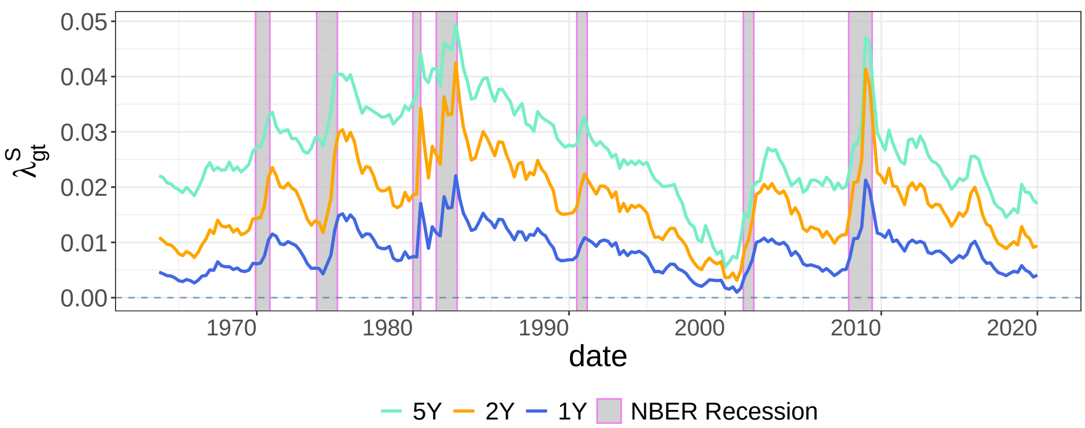
$$\mathbb{E}_t[g_{d,t,t+s}] = \frac{1}{s}\sum_{\tau=1}^{s}\mathbb{E}_t[\Delta d_{t+\tau}], \quad \text{where } \mathbb{E}_t[\Delta d_{t+\tau}] = \mu_g + \sum_{s=\tau}^{\bar{S}}\tilde{\rho}_s f_{t+\tau-s}. \tag{26}$$

Note that in equation (4) the unspanned component w_{gt} is allowed to be persistent. In other words, we do not assume that the MA component of return shocks captures the entire dividend predictability. Consequently, the forward equity yields implied by the MA model are not guaranteed to match the empirical ones exactly.

We present the time series of expected dividend growth and its risk premia for one-, two-, and five-year holding horizons in Figure 8. The estimation is based on a MA(20) formulation, with time-varying risk premia driven by only external predictors. We observe clear business-cycle patterns in the conditional dividend growth and risk premia. While expected dividend growth turns from positive to negative during economic recessions, dividend risk premia generally spike



(a) Expected dividend growth



(b) Dividend risk premia

Figure 8: Time-varying expected dividend growth and risk premia

This figure plots the time-varying expected dividend growth (Panel (a)) and risk premia (Panel (b)). The conditional mean of dividend growth is based on the MA model in equation (26), with $\bar{S} = 20$ quarters. The time-varying dividend risk premia are based on a VAR(1) model for the latent risk factors, in which only the external predictors can forecast latent factors. Estimates are based on the composite cross-section of 275 Fama-French characteristic-sorted portfolios. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

in crisis periods. Furthermore, the term structure of expected dividend growth is downward-sloping in normal times but strongly upward-sloping during recessions; however, we do not detect such patterns for dividend risk premia.

We next estimate forward equity yields for one-, two-, and five-year holding horizons using equation (25), based on the expected dividend growth and risk premia in Figure 8. Note that the estimation is based on the full sample from 1963Q3 to 2019Q4, but we display in Figure 9 the subsample from 2004Q4, the date from which we have the observed data of forward equity yields.[24] Figure 9 shows that our model generates a downward-sloping term structure of equity yields in bad economic states but an upward-sloping one during expansions, consistent with the observed data. Our estimates of forward equity yields are also strongly countercyclical, closely tracking the observed data. Since expected dividend growth displays much more sizable variation than dividend risk premia, the time variation of the latter, rather than the former, explains most of the variation in forward equity yields. This is further highlighted in Figure IA.13, which shows the estimates of forward equity yields based on a (counterfactual) constant risk premia model. Even if dividend risk premia are assumed to be constant, our formulation is able to generate realistic estimates of forward equity yields.

However, we cannot precisely match the observed forward equity yields. For instance, the observed one-year forward equity yield is about 0.3 in 2009Q1, but our model estimates a forward equity yield of only 0.2 in the same period. As we point out in equation (26), our MA formulation allows for other sources of dividend predictability beyond asset returns, which can potentially drive the difference between our estimates and the observed data.

In summary, the term structure of dividend risk premia is different from that of dividend strips, with the gap captured by the expected dividend growth. Our econometric framework targets the term structure of risk premia of not only dividend growth but also other economic quantities. Although our model allows for other sources of dividend predictability beyond the MA representation of priced shocks, we are still able to deliver realistic estimates of forward equity yields, closely matching the observed yields, and generate the time-varying term structures of equity yields.

5 Conclusion

We propose a novel estimator of factors' risk premia, their term structure, and their time variation in a large cross-section of asset returns. The asset returns follow an approximate factor structure, whereas the tested factor can slowly adjust to the asset return systematic shocks, motivated by the Wold decomposition. The latter assumption allows the tested factors and asset returns to have rich dynamics but poses a challenge for the frequentist estimation. We tackle this challenge by taking a Bayesian perspective. Specifically, we derive a Gibbs sampler

[24]The data on realised one-, two-, and five-year forward equity yields are from Bansal, Miller, Song, and Yaron (2021). We thank the authors for sharing the data with us.

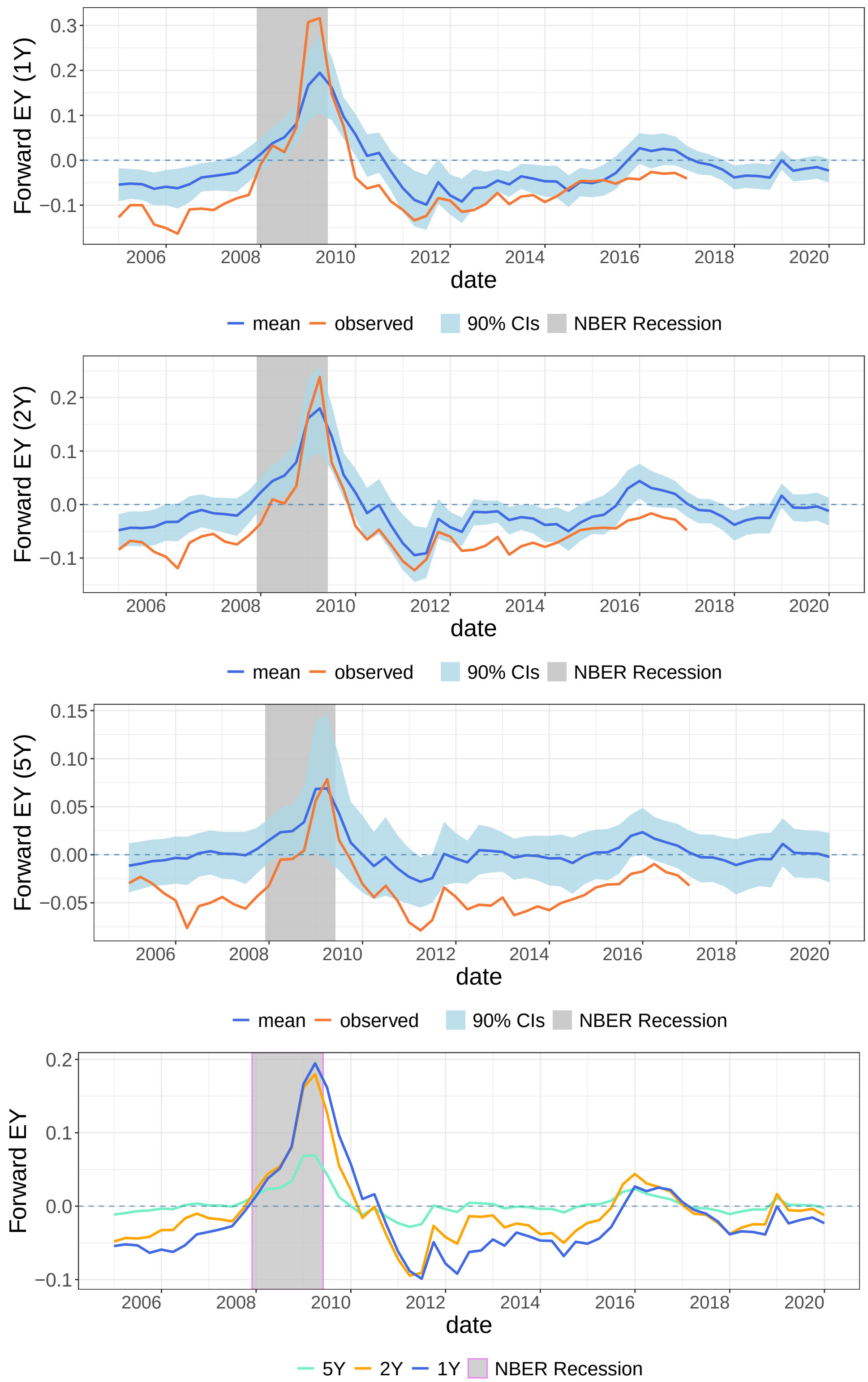


Figure 9: Time series of estimated forward equity yields: Time-varying dividend risk premia

This figure displays the time series of estimated forward equity yields based on our MA model with 20 lags. Dividend risk premia are time-varying and modelled as being linearly dependent on external predictors. We estimate a five-factor latent factor model of FF275 using the full sample from 1963Q3 to 2019Q4. We plot the estimates in the subsample from 2004Q4, the date from which we have the observed data of forward equity yields. The data on realised forward equity yields are from Bansal et al. (2021).

in which all conditional distributions of model parameters have standard closed forms, so our Bayesian estimator is straightforward to implement.

Our Bayesian framework has the frequentist three-pass procedure in Giglio and Xiu (2021) as a particular, unconditional and single-period, case. More precisely, we adopt their rotation invariance property but also show that both the conditional and unconditional term structures of risk premia of observable variables are invariant to arbitrary rotation of the latent factors. We show that the risk premia of an economic state variable over multiple periods can be interpreted as the per-period mean returns of the mimicking portfolios that hedge against its multi-horizon innovations.

We first apply our method to a large equity cross-section. Our results suggest that, unconditionally, most macro variables have significantly upward-sloping term structures of risk premia. Although they are almost unpriced at quarterly horizons, their risk premia, measured over two- to three-year holding horizons, are comparable to many tradable anomalies in equity markets. In other words, macro risk strikes back at business cycle frequencies. Meanwhile, we observe flat or downward-sloping unconditional term structures for other factors, such as VIX and intermediary factors.

Furthermore, conditional on four return predictors often used in previous literature, the macro risk premia are strongly time-varying and have clear business cycle patterns: They are countercyclical, with low risk premia in normal times but significantly increasing risk premia in economic recessions.

Theoretical asset pricing models predict which economic state variables should be priced in the cross-section of asset returns. Given the rich set of new empirical facts we uncover, we argue that when researchers evaluate their models, they should consider the heterogeneous factor risk premia across horizons and states of the business cycle.

References

Adrian, T., E. Etula, and T. Muir (2014). Financial intermediaries and the cross-section of asset returns. *Journal of Finance 69*(6), 2557–2596.

Ai, H., M. M. Croce, A. M. Diercks, and K. Li (2018). News shocks and the production-based term structure of equity returns. *Review of Financial Studies 31*(7), 2423–2467.

Anatolyev, S. and A. Mikusheva (2022). Factor models with many assets: strong factors, weak factors, and the two-pass procedure. *Journal of Econometrics 229*(1), 103–126.

Ang, A., M. Piazzesi, and M. Wei (2006). What does the yield curve tell us about GDP growth? *Journal of Econometrics 131*(1-2), 359–403.

Avramov, D., S. Cheng, L. Metzker, and S. Voigt (2023). Integrating factor models. *Journal of Finance 78*(3), 1593–1646.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Bandi, F. M. and A. Tamoni (2023). Business-cycle consumption risk and asset prices. *Journal of Econometrics*.

Bansal, R., D. Kiku, and A. Yaron (2012). An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review 1*(1), 183–221.

Bansal, R., S. Miller, D. Song, and A. Yaron (2021). The term structure of equity risk premia. *Journal of Financial Economics 142*(3), 1209–1228.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance 59*(4), 1481–1509.

Barillas, F. and J. Shanken (2018). Comparing asset pricing models. *Journal of Finance 73*(2), 715–754.

Belo, F., P. Collin-Dufresne, and R. S. Goldstein (2015). Dividend dynamics and the term structure of dividend strips. *Journal of Finance 70*(3), 1115–1160.

Bryzgalova, S. (2015). Spurious factors in linear asset pricing models. *LSE manuscript 1*(3).

Bryzgalova, S., J. Huang, and C. Julliard (2023). Bayesian solutions for the factor zoo: We just ran two quadrillion models. *Journal of Finance 78*(1), 487–557.

Bryzgalova, S., J. Huang, and C. Julliard (2024). Consumption in asset returns. *Available at SSRN 3783070*.

Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy 107*(2), 205–251.

Campbell, J. Y., S. Giglio, and C. Polk (2013). Hard times. *Review of Asset Pricing Studies 3*(1), 95–132.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies 1*(3), 195–228.

Campbell, J. Y. and T. Vuolteenaho (2004). Bad beta, good beta. *American Economic Review 94*(5), 1249–1275.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance 52*(1), 57–82.

Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica: Journal of the Econometric Society 51*(5), 1281–1304.

Chernov, M., L. A. Lochstoer, and S. R. H. Lundeby (2021). Conditional dynamics and the multihorizon risk-return trade-off. *Review of Financial Studies 35*(3), 1310–1347.

Chib, S., X. Zeng, and L. Zhao (2020). On comparing asset pricing models. *Journal of Finance 75*(1), 551–577.

Cochrane, J. (2009). *Asset Pricing: Revised Edition.* Princeton University Press.

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

Cochrane, J. H. and M. Piazzesi (2008). Decomposing the yield curve. *Available at SSRN 1333274*.

Cohen, R. B., C. Polk, and T. Vuolteenaho (2009). The price is (almost) right. *Journal of Finance 64*(6), 2739–2782.

Connor, G. and R. A. Korajczyk (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics 15*(3), 373–394.

Connor, G. and R. A. Korajczyk (1988). Risk and return in an equilibrium APT: Application of a new test methodology. *Journal of Financial Economics 21*(2), 255–289.

Daniel, K. and D. Marshall (1997). Equity-premium and risk-free-rate puzzles at long horizons. *Macroeconomic Dynamics 1*(2), 452–484.

Dew-Becker, I. and S. Giglio (2016). Asset pricing in the frequency domain: Theory and empirics. *Review of Financial Studies 29*(8), 2029–2068.

Dew-Becker, I., S. Giglio, A. Le, and M. Rodriguez (2017). The price of variance risk. *Journal of Financial Economics 123*(2), 225–250.

Eraker, B. and Y. Wu (2014). Explaining the negative returns to VIX futures and ETNs: An equilibrium approach. *Available at SSRN 2340070*.

Gagliardini, P., E. Ossola, and O. Scaillet (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica 84*(3), 985–1046.

Giglio, S., B. T. Kelly, and S. Kozak (2023). Equity term structures without dividend strips data. *Journal of Finance (forthcoming)*.

Giglio, S. and D. Xiu (2021). Asset pricing with omitted factors. *Journal of Political Economy 129*(7), 1947–1990.

Giglio, S., D. Xiu, and D. Zhang (2023). Test assets and weak factors. NBER Working Paper 29002.

Gospodinov, N., R. Kan, and C. Robotti (2014). Misspecification-robust inference in linear asset-pricing models with irrelevant risk factors. *Review of Financial Studies 27*(7), 2139–2170.

Gospodinov, N., R. Kan, and C. Robotti (2019). Too good to be true? Fallacies in evaluating risk factor models. *Journal of Financial Economics 132*(2), 451–471.

Hansen, L. and R. Jagannathan (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy 99*(2), 225–262.

Hansen, L. P., J. C. Heaton, and N. Li (2008). Consumption strikes back? Measuring long-run risk. *Journal of Political Economy 116*(2), 260–302.

Hasler, M. and R. Marfe (2016). Disaster recovery and the term structure of dividend strips. *Journal of Financial Economics 122*(1), 116–134.

He, Z., B. Kelly, and A. Manela (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics 126*(1), 1–35.

Ingersoll, J. E. (1984). Some results in the theory of arbitrage pricing. *Journal of Finance 39*(4), 1021–1039.

Jagannathan, R. and Y. Wang (2007). Lazy investors, discretionary consumption, and the cross-section of stock returns. *Journal of Finance 62*(4), 1623–1661.

Johnson, T. L. (2017). Risk premia and the VIX term structure. *Journal of Financial and Quantitative Analysis 52*(6), 2461–2490.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review 95*(1), 161–182.

Kan, R., C. Robotti, and J. Shanken (2013). Pricing model performance and the two-pass cross-sectional regression methodology. *Journal of Finance 68*(6), 2617–2649.

Kan, R. and C. Zhang (1999a). GMM tests of stochastic discount factor models with useless factors. *Journal of Financial Economics 54*(1), 103–127.

Kan, R. and C. Zhang (1999b). Two-pass tests of asset pricing models with useless factors. *Journal of Finance 54*(1), 203–235.

Kelly, B., D. Palhares, and S. Pruitt (2023). Modeling corporate bond returns. *Journal of Finance 78*(4), 1967–2008.

Kelly, B. T., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics 134*(3), 501–524.

Kim, D. H. and J. H. Wright (2005). An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates. *Available at SSRN 813267*.

Kleibergen, F. (2009). Tests of risk premia in linear factor models. *Journal of Econometrics 149*(2), 149–173.

Kleibergen, F. and Z. Zhan (2015). Unexplained factors and their effects on second pass R-squared's. *Journal* *of Econometrics 189*(1), 101–116.

Kleibergen, F. and Z. Zhan (2020). Robust inference for consumption-based asset pricing. *Journal of Finance 75*(1), 507–550.

Kozak, S., S. Nagel, and S. Santosh (2018). Interpreting factor models. *Journal of Finance 73*(3), 1183–1223.

Kozak, S., S. Nagel, and S. Santosh (2020). Shrinking the cross-section. *Journal of Financial Economics 135*(2), 271–292.

Kragt, J., F. De Jong, and J. Driessen (2020). The dividend term structure. *Journal of Financial and Quantitative Analysis 55*(3), 829–867.

Lettau, M. and S. Ludvigson (2001). Consumption, aggregate wealth, and expected stock returns. *Journal of Finance 56*(3), 815–849.

Lettau, M., S. C. Ludvigson, and S. Ma (2019). Capital share risk in us asset pricing. *Journal of Finance 74*(4), 1753–1792.

Liew, J. and M. Vassalou (2000). Can book-to-market, size and momentum be risk factors that predict economic growth? *Journal of Financial Economics 57*(2), 221–245.

Lusompa, A. (2023). Local projections, autocorrelation, and efficiency. *Quantitative Economics 14*(4), 1199–1220.

Malloy, C. J., T. J. Moskowitz, and A. Vissing-Jørgensen (2009). Long-run stockholder consumption risk and asset returns. *The Journal of Finance 64*(6), 2427–2479.

McCracken, M. and S. Ng (2020). FRED-QD: A quarterly database for macroeconomic research. Technical report, National Bureau of Economic Research.

Müller, U. K. (2013). Risk of Bayesian inference in misspecified models, and the sandwich covariance matrix. *Econometrica 81*(5), 1805–1849.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Olea, J. L. M., M. Plagborg-Møller, E. Qian, and C. K. Wolf (2024). Double robustness of local projections and some unpleasant varithmetic. Technical report, National Bureau of Economic Research.

Ortu, F., A. Tamoni, and C. Tebaldi (2013). Long-run risk and the persistence of consumption shocks. *Review of Financial Studies 26*(11), 2876–2915.

Parker, J. A. and C. Julliard (2005). Consumption risk and the cross section of expected returns. *Journal of Political Economy 113*(1), 185–222.

Pástor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

van Binsbergen, J. H., M. Brandt, and R. Koijen (2012). On the timing and pricing of dividends. *American Economic Review 102*(4), 1596–1618.

van Binsbergen, J. H. and R. S. Koijen (2017). The term structure of returns: Facts and theory. *Journal of Financial Economics 124*(1), 1–21.

Appendices

Appendix A Additional Propositions and Proofs

Proposition A1. *As $N \to \infty$, $\tilde{\boldsymbol{\mu}}_r^\top \text{cov}(\boldsymbol{r}_t)^{-1}\boldsymbol{\beta}_{\tilde{v}} \to \boldsymbol{\lambda}_{\tilde{v}}^\top$ under the following assumptions:*

i. The K eigenvalues of $\boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\beta}_{\tilde{v}}$ explode as $N \to \infty$, whereas $\boldsymbol{\Sigma}_{wr}$ has bounded eigenvalues: $\gamma(\boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\beta}_{\tilde{v}}) = O_p(N)$ and $\gamma(\boldsymbol{\Sigma}_{wr}) = O_p(1)$;

ii. $\frac{\boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\beta}_{\tilde{v}}}{N}$ and $\boldsymbol{\Sigma}_{wr}$ converge to positive-definite matrices with bounded entries;

iii. Asset returns and their expectations follow equations (1) and (2). In particular, α_i is IID and cross-sectionally independent of factor loadings, with a zero mean and satisfying that $\frac{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}}{N} \to \boldsymbol{0}_K^\top$ as $N \to \infty$. All elements in $\boldsymbol{\beta}_{\tilde{v}}$ are bounded.[25]

A.1 Proof of Proposition A1

Assumptions in equation (2) imply that $\tilde{\boldsymbol{\mu}}_r^\top \text{cov}(\boldsymbol{r}_t)^{-1}\boldsymbol{\beta}_{\tilde{v}} = \underbrace{\boldsymbol{\alpha}^\top \text{cov}(\boldsymbol{r}_t)^{-1}\boldsymbol{\beta}_{\tilde{v}}}_{(I)} + \underbrace{\boldsymbol{\lambda}_{\tilde{v}}^\top \boldsymbol{\beta}_{\tilde{v}}^\top \text{cov}(\boldsymbol{r}_t)^{-1}\boldsymbol{\beta}_{\tilde{v}}}_{(II)}$.

Assumptions in equation (1) imply that $\text{cov}(\boldsymbol{r}_t) = \boldsymbol{\beta}_{\tilde{v}}\boldsymbol{\beta}_{\tilde{v}}^\top + \boldsymbol{\Sigma}_{wr}$. Using the Woodbury matrix identity, we can rewrite the inverse of $\text{cov}(\boldsymbol{r}_t)$ as follows:

$$\text{cov}(\boldsymbol{r}_t)^{-1} = \boldsymbol{\Sigma}_{wr}^{-1} - \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}(\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}})^{-1}\boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}.$$

We now consider the behaviors of components (I) and (II) as $N \to \infty$.

$$\begin{aligned}
(I) &= \boldsymbol{\alpha}^\top \left[\boldsymbol{\Sigma}_{wr}^{-1} - \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}(\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}})^{-1}\boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}\right]\boldsymbol{\beta}_{\tilde{v}} \\
&= \boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}} \cdot \left[\boldsymbol{I}_K - (\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}})^{-1}\boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}\right] \\
&= \boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}} \cdot (\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}})^{-1} = \frac{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}}{N} \cdot \left(\frac{\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}}{N}\right)^{-1}.
\end{aligned}$$

Assumption (ii) in Proposition A.1 implies that $\frac{\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde{v}}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}}{N}$ converges to a positive-definite matrix with bounded entries. On the contrary, due to assumption (iii) in Proposition A.1, $\frac{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}}{N} \to \boldsymbol{0}_K^\top$ as $N \to \infty$, which implies that $(I) \to \boldsymbol{0}_K^\top$.

[25] Ingersoll (1984) defines the pricing errors $\boldsymbol{\alpha}$ such that $\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}} = 0$. Our assumption in Proposition A1 is weaker than that in Ingersoll (1984). This assumption in (iii) is satisfied, e.g., when $\alpha_n \sim O_p(\frac{1}{\sqrt{N}})$, where the latter is a sufficient condition for the absence of asymptotic arbitrage opportunities defined in Ingersoll (1984).

$$
\begin{aligned}
(II) &= \boldsymbol{\lambda}_{\tilde v}^\top \boldsymbol{\beta}_{\tilde v}^\top \left[\boldsymbol{\Sigma}_{wr}^{-1} - \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde v}(\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde v}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde v})^{-1}\boldsymbol{\beta}_{\tilde v}^\top \boldsymbol{\Sigma}_{wr}^{-1}\right]\boldsymbol{\beta}_{\tilde v} \\
&= \boldsymbol{\lambda}_{\tilde v}^\top \left[\boldsymbol{\beta}_{\tilde v}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde v} - \boldsymbol{\beta}_{\tilde v}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde v}(\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde v}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde v})^{-1}\boldsymbol{\beta}_{\tilde v}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde v}\right] \quad (\text{let } \boldsymbol{A} = (\boldsymbol{\beta}_{\tilde v}^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde v})^{-1}) \\
&= \boldsymbol{\lambda}_{\tilde v}^\top \left[\boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}(\boldsymbol{I}_K + \boldsymbol{A}^{-1})^{-1}\boldsymbol{A}^{-1}\right] = \boldsymbol{\lambda}_{\tilde v}^\top (\boldsymbol{A} + \boldsymbol{I}_K)^{-1}.
\end{aligned}
$$

Since we assume that the eigenvalues of $\boldsymbol{\beta}_{\tilde v}^\top \boldsymbol{\beta}_{\tilde v}$ will explode as $N \to \infty$, whereas $\boldsymbol{\Sigma}_{wr}$ has bounded eigenvalues, $\boldsymbol{A} \to \boldsymbol{0}$ as $N \to \infty$. This further implies that $(II) \to \boldsymbol{\lambda}_{\tilde v}^\top$.

A.2 Estimating Time-Varying Risk Premia in Section 2.2

In estimation, we identify a linear rotation of $\tilde{\boldsymbol{v}}_t$: $\boldsymbol{v}_t = \boldsymbol{H}\tilde{\boldsymbol{v}}_t = \boldsymbol{H}\boldsymbol{\mu}_{\tilde v,t-1} + \boldsymbol{H}\boldsymbol{\epsilon}_{\tilde{\boldsymbol{v}}\boldsymbol{t}} = \boldsymbol{\mu}_{v,t-1} + \boldsymbol{\epsilon}_{vt}$, which implies that $\boldsymbol{\Sigma}_{\epsilon v} = \text{cov}(\boldsymbol{\epsilon}_{vt}) = \boldsymbol{H}\boldsymbol{H}^\top$. We generalize the rotation invariance to identify the time-varying risk premia as follows:

$$
\begin{aligned}
&\boldsymbol{r}_t = \boldsymbol{\alpha} - \frac{\boldsymbol{\Upsilon}_r}{2} + \underbrace{\boldsymbol{\beta}_{\tilde v}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\boldsymbol{\lambda}_{\tilde v}}_{\boldsymbol{\lambda}_v} + \underbrace{\boldsymbol{\beta}_{\tilde v}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\tilde{\boldsymbol{v}}_t}_{\boldsymbol{v}_t} + \boldsymbol{w}_{rt}, \quad g_t = \mu_g + \sum_{s=0}^{\bar S}\tilde\rho_s \underbrace{\tilde{\boldsymbol{\eta}}_g^\top \boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_g^\top}\underbrace{\boldsymbol{H}\boldsymbol{\epsilon}_{\tilde v,t-s}}_{\boldsymbol{\epsilon}_{v,t-s}} + w_{gt}, \\
&m_t = 1 - \boldsymbol{\lambda}_v^\top (\boldsymbol{H}^{-1})^\top \boldsymbol{H}^{-1}\boldsymbol{\epsilon}_{vt} - \boldsymbol{\mu}_{v,t-1}^\top (\boldsymbol{H}^{-1})^\top \boldsymbol{H}^{-1}\boldsymbol{\epsilon}_{vt} = 1 - \boldsymbol{\lambda}_v^\top \boldsymbol{\Sigma}_{\epsilon v}^{-1}\boldsymbol{\epsilon}_{vt} - \boldsymbol{\mu}_{v,t-1}^\top \boldsymbol{\Sigma}_{\epsilon v}^{-1}\boldsymbol{\epsilon}_{vt}, \text{ and} \qquad \text{(A1)} \\
&\lambda_{g,t-1}^S = \frac{\sum_{\tau=0}^S \sum_{s=0}^\tau \tilde\rho_s}{1+S} \cdot \underbrace{\tilde{\boldsymbol{\eta}}_g^\top \boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_g^\top}\underbrace{\boldsymbol{H}(\boldsymbol{\lambda}_{\tilde v} + \mathbb{E}_{t-1}\left[\boldsymbol{\mu}_{\tilde v,t+\tau-s-1}\right])}_{\boldsymbol{\lambda}_v + \mathbb{E}_{t-1}\left[\boldsymbol{\mu}_{v,t+\tau-s-1}\right]};
\end{aligned}
$$

therefore, the time-varying risk premia, $\lambda_{g,t-1}^S$, are point identified.

Proposition A2 (Gibbs sampler of the time-varying model)**.** *Under the assumptions in equations (18)–(22), the posterior distribution of the model parameters can be sampled from the following conditional distributions:*

(1) Conditional on the data, $\{g_t\}_{t=1+\bar S}^T$, and shocks to latent factors, $\{\boldsymbol{\epsilon}_{vt}\}_{t=1}^T$, the parameters of the g_t process (σ_{wg}^2, $\boldsymbol{\rho}_g$, and $\boldsymbol{\eta}_g$) follow the normal-inverse-gamma distribution in equations (IA.1)–(IA.3) of Internet Appendix IA.1.1. The only difference is that we replace $\boldsymbol{v}_t$ with $\boldsymbol{\epsilon}_{vt}$ in equations (IA.1)–(IA.3). For point identification purposes, draws of $\boldsymbol{\rho}_g$ and $\boldsymbol{\eta}_g$ are normalized such that $\boldsymbol{\eta}_g^\top \boldsymbol{\eta}_g = 1$.

(2) Conditional on asset returns, $\{\boldsymbol{r}_t\}_{t=1}^T$, and latent factors, $\{\boldsymbol{v}_t\}_{t=1}^T$, the parameters of the $\boldsymbol{r}_t$ process ($\boldsymbol{\Sigma}_{wr}$ and $\boldsymbol{B}_r^\top = (\boldsymbol{\mu}_r, \boldsymbol{\beta}_v)$) follow the normal-inverse-Wishart distribution in equations (IA.4)–(IA.5) of Internet Appendix IA.1.1.

(3) Conditional on asset returns and $(\boldsymbol{\mu}_r, \boldsymbol{\mu}_v, \boldsymbol{\beta}_v, \boldsymbol{\Sigma}_{wr})$, the latent factors, $\boldsymbol{v}_t$, can be sampled from the normal-inverse-Wishart distribution in equation (IA.6).

(4) Conditional on latent factors, $\{\boldsymbol{v}_t\}_{t=1}^T$, the model parameters in the VAR(q) system of $\boldsymbol{v}_t$ can be obtained from equations (IA.9)–(IA.10). The conditional mean of $\boldsymbol{v}_t$ equals the first K elements of $\boldsymbol{\phi}_0 + \boldsymbol{\phi}_1 \boldsymbol{x}_{t-1} + \cdots + \boldsymbol{\phi}_q \boldsymbol{x}_{t-q}$, and the first K variables in $\boldsymbol{\epsilon}_{xt}$ are shocks to priced systematic factors, $\boldsymbol{\epsilon}_{vt}$. We can also obtain the unconditional mean of $\boldsymbol{v}_t$ as the first K elements in $(\boldsymbol{I} - \boldsymbol{\phi}_1 - \cdots - \boldsymbol{\phi}_q)^{-1}\boldsymbol{\phi}_0$.

(5) Conditional on the posterior draws from the time series steps (1)–(4), the posterior distribution of $\boldsymbol{\lambda}_v$ is a Dirac distribution at $(\boldsymbol{\beta}_v^\top \boldsymbol{\beta}_v)^{-1}\boldsymbol{\beta}_v^\top \tilde{\boldsymbol{\mu}}_r$, where $\tilde{\boldsymbol{\mu}}_r = \boldsymbol{\mu}_r + \frac{1}{2}\boldsymbol{\Upsilon}_r$, and $\boldsymbol{\Upsilon}_{ir} = (\boldsymbol{\beta}_v \boldsymbol{\Sigma}_{\epsilon v} \boldsymbol{\beta}_v^\top + \boldsymbol{\Sigma}_{\boldsymbol{wr}})_{ii}$, $i = 1, \ldots, N$. It further yields a Dirac conditional posterior for the term structure of g_t's risk premia at $\lambda_{g,t-1}^S = \sum_{\tau=0}^{S} \sum_{s=0}^{\tau} \frac{\rho_s \boldsymbol{\eta}_g^\top (\boldsymbol{\lambda}_v + \mathbb{E}_{t-1}\left[\boldsymbol{\mu}_{\tilde{v},t+\tau-s-1}\right])}{1+S}$, where $0 \leq S \leq \bar{S}$.

Appendix B Additional Figures and Tables

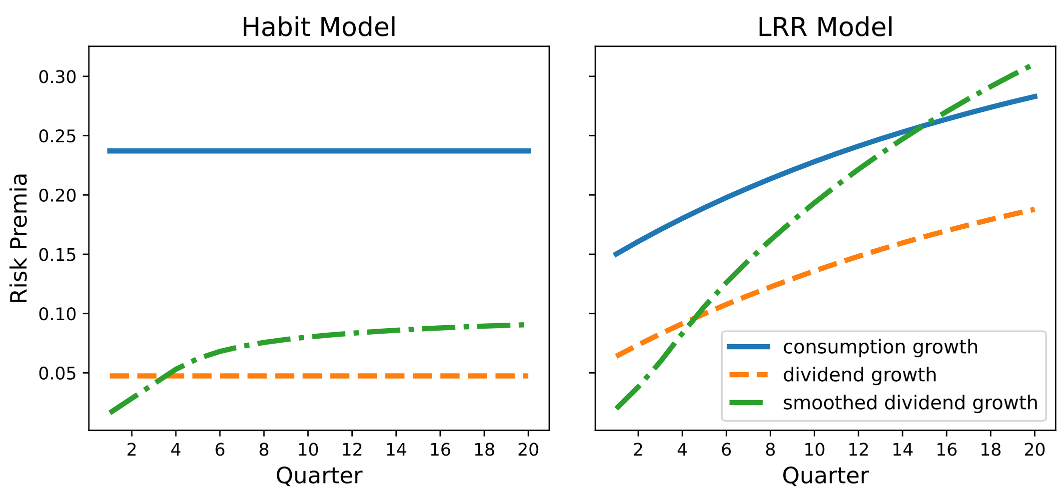


Figure A1: Term Structure of Risk Premia in Habit and Long-Run Risk Models

The figure plots the term structure of risk premia implied by the habit model of Campbell and Cochrane (1999) (left panel) and the long-run risk model (right panel) of Bansal and Yaron (2004). We consider three macro variables: (1) quarterly consumption growth, (2) quarterly dividend growth, and (3) quarterly growth in the smooth dividend payment (the aggregate dividend payments made in the previous 12 months). Risk premia are normalized by the quarterly volatility of the macro variables. Calibration and derivation details can be found in Internet Appendix IA.4.

Table A1: Testing risk premia of strong factors at quarterly frequencies ($T = 200$)

	$S=0$	1	2	3	4	5	6	7	8	9	10	11	12
Panel A: $R_g^2 = 30\%$													
Number of Factors = 5													
10%	0.134	0.110	0.114	0.111	0.113	0.108	0.109	0.107	0.109	0.112	0.110	0.113	0.113
5%	0.075	0.068	0.064	0.068	0.063	0.066	0.064	0.064	0.064	0.066	0.065	0.060	0.063
1%	0.014	0.020	0.019	0.019	0.019	0.016	0.015	0.014	0.012	0.014	0.014	0.014	0.016
Number of Factors = 4													
10%	0.338	0.331	0.331	0.336	0.327	0.328	0.328	0.328	0.325	0.327	0.331	0.326	0.328
5%	0.233	0.225	0.228	0.229	0.233	0.235	0.233	0.232	0.233	0.233	0.219	0.235	0.224
1%	0.089	0.095	0.096	0.094	0.094	0.091	0.090	0.092	0.091	0.091	0.088	0.087	0.089
Number of Factors = 7													
10%	0.141	0.108	0.115	0.124	0.120	0.119	0.111	0.111	0.112	0.117	0.115	0.118	0.119
5%	0.080	0.075	0.074	0.072	0.073	0.069	0.069	0.066	0.071	0.076	0.070	0.071	0.075
1%	0.014	0.018	0.014	0.016	0.017	0.017	0.016	0.017	0.014	0.014	0.015	0.013	0.011
Panel B: $R_g^2 = 20\%$													
Number of Factors = 5													
10%	0.134	0.126	0.122	0.120	0.116	0.115	0.114	0.115	0.115	0.118	0.118	0.123	0.120
5%	0.069	0.064	0.069	0.060	0.059	0.057	0.058	0.058	0.059	0.059	0.057	0.050	0.052
1%	0.008	0.015	0.015	0.013	0.010	0.010	0.011	0.009	0.009	0.011	0.009	0.011	0.012
Number of Factors = 4													
10%	0.307	0.339	0.327	0.320	0.328	0.322	0.327	0.328	0.331	0.336	0.330	0.339	0.337
5%	0.198	0.217	0.219	0.221	0.225	0.220	0.221	0.221	0.218	0.219	0.212	0.214	0.218
1%	0.047	0.085	0.077	0.073	0.078	0.077	0.073	0.077	0.077	0.072	0.067	0.072	0.071
Number of Factors = 7													
10%	0.141	0.131	0.138	0.125	0.128	0.121	0.120	0.125	0.132	0.140	0.142	0.134	0.138
5%	0.074	0.065	0.069	0.066	0.064	0.063	0.067	0.063	0.062	0.057	0.060	0.064	0.062
1%	0.009	0.017	0.013	0.014	0.010	0.010	0.010	0.008	0.011	0.009	0.009	0.012	0.011
Panel C: $R_g^2 = 10\%$													
Number of Factors = 5													
10%	0.117	0.166	0.169	0.159	0.172	0.175	0.174	0.175	0.174	0.173	0.166	0.169	0.172
5%	0.049	0.082	0.085	0.090	0.102	0.099	0.098	0.096	0.095	0.099	0.092	0.091	0.089
1%	0.007	0.018	0.021	0.017	0.024	0.022	0.025	0.029	0.028	0.027	0.019	0.024	0.020
Number of Factors = 4													
10%	0.194	0.287	0.296	0.306	0.313	0.308	0.316	0.318	0.308	0.302	0.284	0.290	0.297
5%	0.093	0.176	0.174	0.173	0.193	0.202	0.188	0.193	0.182	0.187	0.182	0.188	0.184
1%	0.012	0.058	0.055	0.052	0.062	0.061	0.064	0.062	0.066	0.064	0.063	0.063	0.057
Number of Factors = 7													
10%	0.117	0.178	0.168	0.178	0.193	0.197	0.193	0.189	0.191	0.187	0.192	0.186	0.185
5%	0.041	0.100	0.103	0.097	0.112	0.116	0.113	0.113	0.115	0.106	0.104	0.111	0.098
1%	0.004	0.022	0.019	0.017	0.026	0.026	0.025	0.026	0.030	0.031	0.025	0.028	0.023

The table reports the frequency of rejecting the null hypothesis $H_0 : \lambda_g^S = \lambda_g^{S,\star}$ based on the 90%, 95%, and 99% credible intervals of our Bayesian estimates in Proposition 1. λ_g^S is defined in equation (6), and $\lambda_g^{S,\star}$ is λ_g^S's pseudo-true value. We consider strong factors, with $R_g^2 \in \{10\%, 20\%, 30\%\}$. We simulate quarterly observations of g_t and $\boldsymbol{r}_t$ by assuming that i) the true number of latent factors is 5, ii) the time series sample size is 200 quarters, and iii) the true $\bar{S} = 8$. We estimate several model configurations with different numbers of factors (4, 5, and 7) and $\bar{S} = 12$. The number of Monte Carlo simulations is 1,000.

Internet Appendix for:

Macro Strikes Back: Term Structure of Risk Premia

Abstract

The Internet Appendix provides additional propositions, proofs, tables, figures, and empirical results supporting the main text.

IA.1 Additional Propositions and Proofs

IA.1.1 Derivations of the Posterior Distributions in Proposition 1

We present a detailed version of Proposition 1 in the main text.

Proposition IA.1 (Gibbs sampler of the baseline model)**.** *Under the assumptions described in equations (8)–(12), the posterior distribution of model parameters is given by the following conditional distributions:*

(1) Conditional on the data $\{g_t\}_{t=1+\bar{S}}^{T}$ and latent factors $\{\boldsymbol{v}_t\}_{t=1}^{T}$, parameters in g_t's equation follow a normal-inverse-gamma distribution:

$$\sigma_{wg}^2 \mid \{g_t\}_{t=1+\bar{S}}^{T}, \boldsymbol{\rho}_g, \boldsymbol{\eta}_g, \{\boldsymbol{v}_t\}_{t=1}^{T} \sim \mathcal{IG}\left(\frac{T-\bar{S}}{2}, \frac{(\boldsymbol{G}-\boldsymbol{V}_{\rho}\boldsymbol{\rho}_g)^\top(\boldsymbol{G}-\boldsymbol{V}_{\rho}\boldsymbol{\rho}_g)}{2}\right), \tag{IA.1}$$

$$\boldsymbol{\rho}_g \mid \boldsymbol{G}, \sigma_{wg}^2, \boldsymbol{\eta}_g, \{\boldsymbol{v}_t\}_{t=1}^{T} \sim \mathcal{N}\left(\left(\boldsymbol{V}_{\boldsymbol{\rho}}^\top \boldsymbol{V}_{\boldsymbol{\rho}}\right)^{-1} \boldsymbol{V}_{\boldsymbol{\rho}}^\top \boldsymbol{G},\ \hat{\boldsymbol{\Sigma}}_{\rho}\right), \text{ and} \tag{IA.2}$$

$$\boldsymbol{\eta}_g \mid \boldsymbol{G}, \sigma_{wg}^2, \boldsymbol{\rho}_g, \{\boldsymbol{v}_t\}_{t=1}^{T} \sim \mathcal{N}\left(\left(\boldsymbol{V}_{\boldsymbol{\eta}}^\top \boldsymbol{V}_{\boldsymbol{\eta}}\right)^{-1} \boldsymbol{V}_{\boldsymbol{\eta}}^\top \bar{\boldsymbol{G}},\ \hat{\boldsymbol{\Sigma}}_{\eta}\right). \tag{IA.3}$$

To identify $\boldsymbol{\rho}_g$ and $\boldsymbol{\eta}_g$, we normalize $\boldsymbol{\eta}_g$ after each posterior draw such that $\boldsymbol{\eta}_g^\top \boldsymbol{\eta}_g = 1$.

(2) Conditional on asset returns and latent factors, we update model parameters in $\boldsymbol{r}_t$'s equation using a normal-inverse-Wishart distribution, as follows:

$$\boldsymbol{\Sigma}_{wr} \mid \boldsymbol{R}, \{\boldsymbol{v}_t\}_{t=1}^{T}, \boldsymbol{\mu}_r, \boldsymbol{\beta}_v \sim \mathcal{W}^{-1}\left(T,\ \left(\boldsymbol{R}-\boldsymbol{V}_r\boldsymbol{B}_r\right)^\top\left(\boldsymbol{R}-\boldsymbol{V}_r\boldsymbol{B}_r\right)\right) \text{ and} \tag{IA.4}$$

$$\boldsymbol{B}_r \mid \boldsymbol{R}, \{\boldsymbol{v}_t\}_{t=1}^{T}, \boldsymbol{\Sigma}_{wr} \sim \mathcal{MVN}\left((\boldsymbol{V}_r^\top \boldsymbol{V}_r)^{-1}\boldsymbol{V}_r^\top \boldsymbol{R},\ \boldsymbol{\Sigma}_{wr} \otimes (\boldsymbol{V}_r^\top \boldsymbol{V}_r)^{-1}\right), \tag{IA.5}$$

where $\boldsymbol{B}_r^\top = (\boldsymbol{\mu}_r, \boldsymbol{\beta}_v)$.

(3) Conditional on asset returns and $(\boldsymbol{\mu}_r, \boldsymbol{\beta}_v, \boldsymbol{\Sigma}_{wr})$, we update both latent factors $\boldsymbol{v}_t$ and their mean and covariance parameters, as follows:

$$\boldsymbol{v}_t \mid \boldsymbol{r}_t, \boldsymbol{\mu}_r, \boldsymbol{\beta}_v, \boldsymbol{\Sigma}_{wr}, \boldsymbol{\mu}_v, \boldsymbol{\Sigma}_v \sim \mathcal{N}\left(\left(\boldsymbol{\beta}_{\boldsymbol{v}}^\top \boldsymbol{\Sigma}_{\boldsymbol{wr}}^{-1}\boldsymbol{\beta}_{\boldsymbol{v}}\right)^{-1}\left[\boldsymbol{\beta}_{\boldsymbol{v}}^\top \boldsymbol{\Sigma}_{\boldsymbol{wr}}^{-1}(\boldsymbol{r}_t - \boldsymbol{\mu}_r + \boldsymbol{\beta}_{\boldsymbol{v}}\boldsymbol{\mu}_v)\right],\ \left(\boldsymbol{\beta}_{\boldsymbol{v}}^\top \boldsymbol{\Sigma}_{\boldsymbol{wr}}^{-1}\boldsymbol{\beta}_{\boldsymbol{v}}\right)^{-1}\right), \tag{IA.6}$$

$$\boldsymbol{\Sigma}_v \mid \{\boldsymbol{v}_t\}_{t=1}^T \sim \mathcal{W}^{-1}\Big(T-1,\ \sum_{t=1}^T (\boldsymbol{v}_t - \bar{\boldsymbol{v}})(\boldsymbol{v}_t - \bar{\boldsymbol{v}})^\top\Big),\ \textit{and} \tag{IA.7}$$

$$\boldsymbol{\mu}_v \mid \boldsymbol{\Sigma}_v, \{\boldsymbol{v}_t\}_{t=1}^T \sim \mathcal{N}\left(\bar{\boldsymbol{v}},\ \boldsymbol{\Sigma}_v/T\right), \tag{IA.8}$$

where $\bar{\boldsymbol{v}} = \sum_{t=1}^T \boldsymbol{v}_t/T$. *In steps (1)–(3),* $\mathcal{IG}(\cdot)$ *denotes the inverse-gamma distribution,* $\mathcal{N}(\cdot)$ *and* $\mathcal{MVN}(\cdot)$ *denote the normal and multivariate normal distributions, and* $\mathcal{W}^{-1}(\cdot)$ *is the inverse-Wishart distribution. The quantities* $\boldsymbol{G}$, $\bar{\boldsymbol{G}}$, $\boldsymbol{V}_\rho$, $\boldsymbol{V}_\eta$, $\hat{\boldsymbol{\Sigma}}_\rho$, $\hat{\boldsymbol{\Sigma}}_\eta$, $\boldsymbol{V}_r$, *and* $\boldsymbol{R}$ *are defined in the proof.*

(4) Based on the posterior draws from the time series steps (1)–(3), the posterior distribution of $\boldsymbol{\lambda}_v$ *is a Dirac distribution at* $(\boldsymbol{\beta}_v^\top \boldsymbol{\beta}_v)^{-1}\boldsymbol{\beta}_v^\top \tilde{\boldsymbol{\mu}}_r$. *In addition, the posterior distribution of the term structure of* g_t*'s risk premia is also a Dirac distribution at* $\lambda_g^S = \frac{\sum_{\tau=0}^S \sum_{s=0}^\tau \rho_s}{1+S} \cdot \boldsymbol{\eta}_g^\top \boldsymbol{\lambda}_{\boldsymbol{v}}$, *where* $0 \le S \le \bar{S}$.

We next derive the posterior distribution in g_t's equation. We introduce some matrix notations, as follows:

$$\boldsymbol{V}_\rho = \begin{pmatrix} 1 & (\boldsymbol{v}_{\bar{S}+1} - \boldsymbol{\mu}_v)^\top \boldsymbol{\eta}_g & \cdots & (\boldsymbol{v}_1 - \boldsymbol{\mu}_v)^\top \boldsymbol{\eta}_g \\ \vdots & \vdots & & \vdots \\ 1 & (\boldsymbol{v}_T - \boldsymbol{\mu}_v)^\top \boldsymbol{\eta}_g & \cdots & (\boldsymbol{v}_{T-\bar{S}} - \boldsymbol{\mu}_v)^\top \boldsymbol{\eta}_g \end{pmatrix},$$

$$\boldsymbol{V}_\eta = \begin{pmatrix} \sum_{s=0}^{\bar{S}} \rho_s (\boldsymbol{v}_{1,1+\bar{S}-s} - \boldsymbol{\mu}_v) & \cdots & \sum_{s=0}^{\bar{S}} \rho_s (\boldsymbol{v}_{K,1+\bar{S}-s} - \boldsymbol{\mu}_v) \\ \vdots & & \vdots \\ \sum_{s=0}^{\bar{S}} \rho_s (\boldsymbol{v}_{1,T-s} - \boldsymbol{\mu}_v) & \cdots & \sum_{s=0}^{\bar{S}} \rho_s (\boldsymbol{v}_{K,T-s} - \boldsymbol{\mu}_v) \end{pmatrix},$$

$$\boldsymbol{G} = (g_{1+\bar{S}}, \ldots, g_T)^\top, \quad \text{and } \bar{\boldsymbol{G}} = (g_{1+\bar{S}} - \mu_g, \ldots, g_T - \mu_g)^\top.$$

Using the notations above, the data likelihood for $\boldsymbol{G}$ can be written as

$$p(\boldsymbol{G} \mid \boldsymbol{\rho}_g, \boldsymbol{\eta}_g, \{\boldsymbol{v}_t\}_{t=1}^T, \sigma_{wg}^2) = (2\pi\sigma_{wg}^2)^{-\frac{T-\bar{S}}{2}} \exp\Big\{-\frac{1}{2\sigma_{wg}^2}(\boldsymbol{G} - \boldsymbol{V}_\rho\boldsymbol{\rho}_g)^\top(\boldsymbol{G} - \boldsymbol{V}_\rho\boldsymbol{\rho}_g)\Big\},$$

where $(\boldsymbol{G} - \boldsymbol{V}_\rho\boldsymbol{\rho}_g)^\top(\boldsymbol{G} - \boldsymbol{V}_\rho\boldsymbol{\rho}_g) = (\bar{\boldsymbol{G}} - \boldsymbol{V}_\eta\boldsymbol{\eta}_g)^\top(\bar{\boldsymbol{G}} - \boldsymbol{V}_\eta\boldsymbol{\eta}_g)$. Since we assign a flat prior to $(\boldsymbol{\rho}_g, \boldsymbol{\eta}_g, \sigma_{wg}^2)$, the posterior distribution of σ_{wg}^2 is

$$p(\sigma_{wg}^2 \mid \boldsymbol{G}, \boldsymbol{\rho}_g, \boldsymbol{\eta}_g, \{\boldsymbol{v}_t\}_{t=1}^T) \propto \left(\frac{1}{\sigma_{wg}^2}\right)^{\frac{T-\bar{S}}{2}+1} \exp\Big\{-\frac{(\boldsymbol{G} - \boldsymbol{V}_\rho\boldsymbol{\rho}_g)^\top(\boldsymbol{G} - \boldsymbol{V}_\rho\boldsymbol{\rho}_g)}{2\sigma_{wg}^2}\Big\};$$

hence, the posterior distribution of σ_{wg}^2 is an inverse-gamma in equation (IA.1).

We next consider the posterior distribution of $\boldsymbol{\rho}_g$ and $\boldsymbol{\eta}_g$. From the data likelihood, we can derive the kernel of $\boldsymbol{\rho}_g$'s posterior,

$$p(\boldsymbol{\rho}_g \mid \boldsymbol{G}, \sigma^2_{wg}, \boldsymbol{\eta}_g, \{\boldsymbol{v}_t\}_{t=1}^T) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\rho}_g - \hat{\boldsymbol{\rho}}_g)^\top \left[\sigma^2_{wg}\big(\boldsymbol{V}_{\boldsymbol{\rho}}^\top \boldsymbol{V}_{\boldsymbol{\rho}}\big)^{-1}\right]^{-1}(\boldsymbol{\rho}_g - \hat{\boldsymbol{\rho}}_g)\right\},$$

where $\hat{\boldsymbol{\rho}}_g = \big(\boldsymbol{V}_{\boldsymbol{\rho}}^\top \boldsymbol{V}_{\boldsymbol{\rho}}\big)^{-1}\boldsymbol{V}_{\boldsymbol{\rho}}^\top \boldsymbol{G}$. The next step is to make adjustments for the posterior covariance matrix of $\boldsymbol{\rho}_g$ due to the potentially autocorrelated $w_{gt}\boldsymbol{V}_{\rho t}$. A simple solution is given by Müller (2013), which proposes that we can replace $\sigma^2_{wg}\big(\boldsymbol{V}_{\boldsymbol{\rho}}^\top \boldsymbol{V}_{\boldsymbol{\rho}}\big)^{-1}$ with the Newey and West (1987) type of sandwich covariance matrix, denoted as $\hat{\boldsymbol{\Sigma}}_\rho$, as follows:

$$\boldsymbol{\rho}_g \mid \boldsymbol{G}, \sigma^2_{wg}, \boldsymbol{\eta}_g, \{\boldsymbol{v}_t\}_{t=1}^T \sim \mathcal{N}\left(\hat{\boldsymbol{\rho}}_g, \hat{\boldsymbol{\Sigma}}_\rho\right), \quad \hat{\boldsymbol{\Sigma}}_\rho = \big(\boldsymbol{V}_{\boldsymbol{\rho}}^\top \boldsymbol{V}_{\boldsymbol{\rho}}\big)^{-1}\big[(T-\bar{S})\hat{\boldsymbol{S}}_\rho\big]\big(\boldsymbol{V}_{\boldsymbol{\rho}}^\top \boldsymbol{V}_{\boldsymbol{\rho}}\big)^{-1},$$

$$\hat{\boldsymbol{S}}_\rho = \frac{1}{T-\bar{S}}\sum_{t=1+\bar{S}}^{T} \hat{w}^2_{g,t}\big(\boldsymbol{V}_{\rho,t}\boldsymbol{V}_{\rho,t}^\top\big) + \sum_{l=1}^{L}\big(1-\frac{l}{1+L}\big)\hat{\boldsymbol{\Gamma}}_{\rho l}, \text{ and}$$

$$\hat{\boldsymbol{\Gamma}}_{\rho l} = \frac{1}{T-\bar{S}-l}\sum_{t=1+\bar{S}+l}^{T} \hat{w}_{g,t}\hat{w}_{g,t-l}\big(\boldsymbol{V}_{\rho,t}\boldsymbol{V}_{\rho,t-l}^\top + \boldsymbol{V}_{\rho,t-l}\boldsymbol{V}_{\rho,t}^\top\big) \text{ for } l>0, \quad \hat{w}_{g,t} = g_t - \boldsymbol{V}_{\rho t}^\top \hat{\boldsymbol{\rho}}_g,$$

where L, the number of lags in the Newey-West estimator, is chosen to be $\bar{S}$ since $w_{gt}\boldsymbol{V}_{\rho t}$ and $w_{g,t-l}\boldsymbol{V}_{\rho t-l}$ are uncorrelated for $l > \bar{S}$.

We finish deriving the multivariate normal in equation (IA.2). A similar derivation can be applied to the posterior distribution of $\boldsymbol{\eta}_g$ in equation (IA.3).

We now proceed to derive the posterior distribution of model parameters in $\boldsymbol{r}_t$'s equation. We stack time series observations into the following matrices:

$$\boldsymbol{R} = \begin{pmatrix} \boldsymbol{r}_1^\top \\ \vdots \\ \boldsymbol{r}_T^\top \end{pmatrix}, \quad \boldsymbol{V}_r = \begin{pmatrix} 1 & (\boldsymbol{v}_1 - \boldsymbol{\mu}_v)^\top \\ \vdots & \vdots \\ 1 & (\boldsymbol{v}_T - \boldsymbol{\mu}_v)^\top \end{pmatrix}, \quad \text{and } \boldsymbol{B}_r = \begin{pmatrix} \boldsymbol{\mu}_r^\top \\ \boldsymbol{\beta}_v^\top \end{pmatrix},$$

and the data likelihood of asset returns is

$$p(\boldsymbol{R} \mid \{\boldsymbol{v}_t\}_{t=1}^T, \boldsymbol{\mu}_r, \boldsymbol{\beta}_v, \boldsymbol{\Sigma}_{wr}) \propto |\boldsymbol{\Sigma}_{wr}|^{-\frac{T}{2}}\exp\big\{-\frac{1}{2}tr\big[\boldsymbol{\Sigma}_{wr}^{-1}\big(\boldsymbol{R}-\boldsymbol{V}_r\boldsymbol{B}_r\big)^\top\big(\boldsymbol{R}-\boldsymbol{V}_r\boldsymbol{B}_r\big)\big]\big\}.$$

Under the prior distribution in equation (11), we first derive the posterior of $\boldsymbol{\Sigma}_{wr}$,

$$p(\boldsymbol{\Sigma}_{wr} \mid \boldsymbol{R}, \{\boldsymbol{v}_t\}_{t=1}^T, \boldsymbol{\mu}_r, \boldsymbol{\beta}_v) \propto |\boldsymbol{\Sigma}_{wr}|^{-\frac{T+N+1}{2}}\exp\big\{-\frac{1}{2}tr\big[\boldsymbol{\Sigma}_{wr}^{-1}\big(\boldsymbol{R}-\boldsymbol{V}_r\boldsymbol{B}_r\big)^\top\big(\boldsymbol{R}-\boldsymbol{V}_r\boldsymbol{B}_r\big)\big]\big\},$$

which implies the inverse-Wishart distribution of $\boldsymbol{\Sigma}_{wr}$ in equation (IA.4). When $\boldsymbol{\Sigma}_{wr}$ is diagonal, which is assumed in the high-dimensional setting, the inverse-Wishart distribution reduces to independent inverse-gamma distributions of $\{\sigma^2_{wr,n}\}_{n=1}^{N}$.

We next derive the posterior of $(\boldsymbol{\mu}_r, \boldsymbol{\beta}_v)$

$$p(\boldsymbol{B}_r \mid \boldsymbol{R}, \{\boldsymbol{v}_t\}_{t=1}^T, \boldsymbol{\Sigma}_{wr}) \propto \exp\big\{-\frac{1}{2}tr\big[\boldsymbol{\Sigma}_{wr}^{-1}\big(\boldsymbol{B}_r - \hat{\boldsymbol{B}}_r\big)^\top \boldsymbol{V}_r^\top \boldsymbol{V}_r\big(\boldsymbol{B}_r - \hat{\boldsymbol{B}}_r\big)\big]\big\},$$

where $\hat{\boldsymbol{B}}_r = (\boldsymbol{V}_r^\top \boldsymbol{V}_r)^{-1}\boldsymbol{V}_r^\top \boldsymbol{R}$, and the formula above is the kernel of the multivariate normal distribution in equation (IA.5). However, when we implement equation (IA.5), we replace $(\boldsymbol{V}_r^\top \boldsymbol{V}_r)^{-1}$ with $(\boldsymbol{V}_r^\top \boldsymbol{V}_r + \boldsymbol{D}_r)^{-1}$, where $\boldsymbol{D}_r = \text{diag}\{0, 1, \ldots, 1\}$. The additional term $\boldsymbol{D}_r$ is a small penalty that preempts numerical difficulties in high-dimensional applications.

Finally, we derive the posterior distribution of latent factors and their means and covariance matrix. The posterior distribution of $\boldsymbol{v}_t$ is

$$\begin{aligned}
&p(\boldsymbol{v}_t \mid \boldsymbol{r}_t, \boldsymbol{\mu}_r, \boldsymbol{\beta}_v, \boldsymbol{\Sigma}_{wr}, \boldsymbol{\mu}_v, \boldsymbol{\Sigma}_v) \\
\propto\ & p(\boldsymbol{r}_t \mid \boldsymbol{v}_t, \boldsymbol{\mu}_r, \boldsymbol{\beta}_v, \boldsymbol{\Sigma}_{wr})\pi(\boldsymbol{v}_t \mid \boldsymbol{\mu}_v, \boldsymbol{\Sigma}_v) \\
\propto\ & \exp\big\{-\frac{1}{2}(\boldsymbol{r}_t - \boldsymbol{\mu}_r + \boldsymbol{\beta}_v\boldsymbol{\mu}_v - \boldsymbol{\beta}_v\boldsymbol{v}_t)^\top \boldsymbol{\Sigma}_{wr}^{-1}(\boldsymbol{r}_t - \boldsymbol{\mu}_r + \boldsymbol{\beta}_v\boldsymbol{\mu}_v - \boldsymbol{\beta}_v\boldsymbol{v}_t)\big\} \\
\propto\ & \exp\big\{-\frac{1}{2}\big[\boldsymbol{v}_t^\top(\boldsymbol{\beta}_v^\top \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_v)\boldsymbol{v}_t - 2\boldsymbol{v}_t^\top \boldsymbol{\beta}_v^\top \boldsymbol{\Sigma}_{wr}^{-1}(\boldsymbol{r}_t - \boldsymbol{\mu}_r + \boldsymbol{\beta}_v\boldsymbol{\mu}_v)\big]\big\},
\end{aligned}$$

which implies equation (IA.6). The posterior distribution of $(\boldsymbol{\mu}_v, \boldsymbol{\Sigma}_v)$ is

$$p(\boldsymbol{\mu}_v, \boldsymbol{\Sigma}_v \mid \{\boldsymbol{v}_t\}_{t=1}^T) \propto |\boldsymbol{\Sigma}_v|^{\frac{T+K+1}{2}} \exp\big\{-\frac{1}{2}tr\big[\boldsymbol{\Sigma}_v^{-1}\sum_{t=1}^{T}(\boldsymbol{v}_t - \boldsymbol{\mu}_v)(\boldsymbol{v}_t - \boldsymbol{\mu}_v)^\top\big]\big\},$$

which is the kernel of the normal-inverse-Wishart distribution in equations (IA.7) and (IA.8).

IA.1.2 Proof of Proposition A2

The only new ingredient in Proposition A2 is step 4, which estimates the model parameters in the VAR(q) system of $\boldsymbol{x}_t$. First, we introduce the following matrix notations:

$$\boldsymbol{X}^{(1)} = \begin{pmatrix} \boldsymbol{x}_{q+1}^\top \\ \vdots \\ \boldsymbol{x}_T^\top \end{pmatrix}, \quad \boldsymbol{X}^{(0)} = \begin{pmatrix} 1 & \boldsymbol{x}_q^\top & \ldots & \boldsymbol{x}_1^\top \\ \vdots & \vdots & & \vdots \\ 1 & \boldsymbol{x}_{T-1}^\top & \ldots & \boldsymbol{x}_{T-q}^\top \end{pmatrix}, \quad \text{and } \boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\phi}_0^\top \\ \vdots \\ \boldsymbol{\phi}_q^\top \end{pmatrix},$$

and equation (22) implies that the data likelihood is

$$p(\boldsymbol{X}^{(1)} \mid \boldsymbol{X}^{(0)}, \boldsymbol{\Phi}, \boldsymbol{\Sigma}_{\epsilon x}) \propto |\boldsymbol{\Sigma}_{\epsilon x}|^{-\frac{T-q}{2}} \exp\Big\{-\frac{1}{2} tr\big[\boldsymbol{\Sigma}_{\epsilon x}^{-1}\big(\boldsymbol{X}^{(1)} - \boldsymbol{X}^{(0)}\boldsymbol{\Phi}\big)^\top\big(\boldsymbol{X}^{(1)} - \boldsymbol{X}^{(0)}\boldsymbol{\Phi}\big)\big]\Big\}.$$

Under the prior distribution $\pi(\boldsymbol{\Phi}, \boldsymbol{\Sigma}_{\epsilon x}) \propto |\boldsymbol{\Sigma}_{\epsilon x}|^{-\frac{K+p+1}{2}}$, we can easily show that $(\boldsymbol{\Phi}, \boldsymbol{\Sigma}_{\epsilon x})$ follow the normal-inverse-Wishart distribution,

$$\boldsymbol{\Sigma}_{\epsilon x} \mid \boldsymbol{\Phi}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(0)}, \mathcal{W}^{-1}\Bigg(T-q,\ \big(\boldsymbol{X}^{(1)} - \boldsymbol{X}^{(0)}\boldsymbol{\Phi}\big)^\top\big(\boldsymbol{X}^{(1)} - \boldsymbol{X}^{(0)}\boldsymbol{\Phi}\big)\Bigg) \text{ and} \tag{IA.9}$$

$$\boldsymbol{\Phi} \mid \boldsymbol{\Sigma}_{\epsilon x}, \boldsymbol{X}^{(1)}, \boldsymbol{X}^{(0)} \sim \mathcal{MVN}\Bigg(\big((\boldsymbol{X}^{(0)})^\top \boldsymbol{X}^{(0)}\big)^{-1}(\boldsymbol{X}^{(0)})^\top \boldsymbol{X}^{(1)},\ \boldsymbol{\Sigma}_{\epsilon x} \otimes \big((\boldsymbol{X}^{(0)})^\top \boldsymbol{X}^{(0)}\big)^{-1}\Bigg), \tag{IA.10}$$

following similar derivations as in equations (IA.4)–(IA.5).

IA.2 Simulations: Time-Varying Risk Premia

In this section, we explore the finite-sample performance of our Bayesian estimator in Proposition A2 when latent factors command time-varying risk premia. Different from the unconditional risk premia model, we simulate latent factors, $\tilde{\boldsymbol{v}}_t$, from the VAR(1) process,

$$\tilde{\boldsymbol{v}}_t = \hat{\boldsymbol{\phi}}_1 \tilde{\boldsymbol{v}}_{t-1} + \boldsymbol{\epsilon}_{\tilde{v}t}, \quad \boldsymbol{\epsilon}_{\tilde{v}t} \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}_K, \boldsymbol{I}_K),$$

where $\hat{\boldsymbol{\phi}}_1$ is calibrated by running the VAR(1) regression using the top five PCs of asset returns.[1] Next, we simulate the asset returns as before, assuming a five-factor model. Finally, we generate g_t such that it is driven by $\boldsymbol{\epsilon}_{\tilde{v}t}$ instead of $\tilde{\boldsymbol{v}}_t$:

$$g_t = c \cdot \sum_{s=0}^{\bar{S}} \hat{\rho}_s f_{t-s} + w_{gt}, \quad f_t = \frac{1}{\sqrt{3}}(1, 0, 1, 0, 1)\boldsymbol{\epsilon}_{\tilde{v}t}.$$

Similar to the simulations in the main text, we report the size, power, and time series fit in g_t's equation (R_g^2) and the correlation between estimated and pseudo-true latent process f_t in Tables IA.IX–IA.VIII. Overall, our Bayesian estimator in Proposition A2 has satisfactory finite-sample performance, delivering consistent estimates of unconditional risk premia when the priced systematic factors follow a VAR(1) process.

[1] If a parameter in $\boldsymbol{\phi}_1$ is not significant at the 10% level, we set it to be zero in $\hat{\boldsymbol{\phi}}_1$.

IA.3 Data Description

We consider a cross-section of 275 equity portfolios collected from Ken French's website (FF275): 25 (5×5) portfolios sorted by (1) size and book-to-market ratio, (2) size and accrual, (3) size and beta, (4) size and investment, (5) size and long-term reversals, (6) size and momentum, (7) size and net issuance, (8) size and profitability, (9) size and residual variance, (10) size and variance, and (11) size and short-term reversals. The sample ranges from Q3 1963 to Q4 2019.

Table IA.II presents the factors studied in Section 4. We show each variable's name, description, sample, and data source. When the sample of factors differs from that of asset returns, we use the overlapping sample. Hence, different factors use different samples in estimation. We briefly describe how we construct the macro PCs using FRED-QD. There are 246 macro variables in the dataset, but we keep only those with complete observations. We next estimate the correlation structure of the remaining 159 variables and use it to construct the five PCs, which account for 25.5%, 9.4%, 5.3%, 4.7%, and 4.4% of the time series variations of this large panel data.

IA.4 Term Structure of Risk Premia in Macro-Finance Models

The first model that we consider is the external habit model of Campbell and Cochrane (1999) (CC henceforth). The model dynamics are summarized by the following equations:

$$\begin{aligned}
\text{log SDF:} \quad & m_{t+1} = \log\delta - \gamma g + \gamma(1-\phi)(s_t - \bar{s}) - \gamma[1+\lambda(s_t)]v_{t+1}, \\
\text{log consumption surplus ratio:} \quad & s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)v_{t+1}, \\
\text{log consumption growth:} \quad & \Delta c_{t+1} = g + v_{t+1}, \quad v_{t+1} \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma^2), \text{ and} \\
\text{log dividend growth:} \quad & \Delta d_{t+1} = g + w_{t+1}, \quad w_{t+1} \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma_w^2), \quad \text{corr}(w_t, v_t) = \rho,
\end{aligned}$$

where v_t and w_t are shocks to consumption and dividend growth, respectively, and their correlation equals ρ. CC choose the specification of $\lambda(s_t)$ to ensure a constant risk-free rate:

$$\lambda(s_t) = \begin{cases} \frac{\sqrt{1-2(s_t-\bar{s})}}{\bar{S}} - 1, & s_t \leq s_{max} \\ 0, & s_t > s_{max} \end{cases}, \quad \text{where: } \bar{S} = \sigma\sqrt{\frac{\gamma}{1-\phi}}, \quad s_{max} = \bar{s} + \frac{1}{2}(1-\bar{S}^2).$$

We simulate the dynamics of $(m_{t+1}, s_{t+1}, \Delta c_{t+1}, \Delta d_{t+1}, \lambda(s_t))$ following the same parameter choices as in Table 1 of CC, which are summarized in Panel A of Table IA.I.

Secondly, we consider the long-run risk model of Bansal and Yaron (2004) (BY henceforth), in which they introduce slow-moving conditional mean and stochastic volatility of consumption and dividend growth. We summarize the dynamics of the state variables as follows:

$$\begin{aligned}
\text{conditional consumption mean:} \quad & x_{t+1} = \rho x_t + \varphi_e \sigma_t e_{t+1} \\
\text{log consumption growth:} \quad & \Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{t+1} \\
\text{log dividend growth:} \quad & \Delta d_{t+1} = \mu_d + \phi_d x_t + \pi \sigma_t \eta_{t+1} + \varphi_d \sigma_t u_{t+1}, \text{ and} \\
\text{stochastic volatility:} \quad & \sigma_{t+1}^2 = \sigma^2 + \nu_1(\sigma_t^2 - \sigma^2) + \sigma_\omega \omega_{t+1},
\end{aligned}$$

where $e_{t+1},\ u_{t+1},\ \eta_{t+1}, \omega_{t+1} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$.

To solve the model, BY consider the approximate solution for the price-consumption ratio, that is, $z_t = A_0 + A_1 x_t + A_2 \sigma_t^2$, where

$$A_0 = \frac{1}{1-\kappa_1}\left[\log\delta + \kappa_0 + (1-\frac{1}{\psi})\mu + \kappa_1 A_2(1-\nu_1)\sigma^2 + \frac{\theta}{2}(\kappa_1 A_2 \sigma_\omega)^2\right], \quad A_1 = \frac{1-\frac{1}{\psi}}{1-\kappa_1\rho},$$

$$A_2 = \frac{0.5\cdot\left[\left(\theta-\frac{\theta}{\psi}\right)^2 + (\theta A_1\kappa_1\varphi_e)^2\right]}{\theta(1-\kappa_1\nu_1)}, \quad \kappa_1 = \frac{\exp(\bar{z})}{1+\exp(\bar{z})}, \quad \text{and } \kappa_0 = \log\left(1+\exp(\bar{z})\right) - \kappa_1\bar{z}.$$

The steady state $\bar{z}$ can be found by numerically solving a fixed-point problem: $\bar{z} = A_0(\bar{z}) + A_2(\bar{z})\sigma^2$. Finally, the shock in the log SDF is

$$m_{t+1} - \mathbb{E}_t(m_{t+1}) = \lambda_{m,\eta}\sigma_t\eta_{t+1} - \lambda_{m,e}\sigma_t e_{t+1} - \lambda_{m,\omega}\sigma_\omega\omega_{t+1},$$

where $\lambda_{m,\eta} = \left[-\frac{\theta}{\psi} + \theta - 1\right] = -\gamma, \quad \lambda_{m,e} = (1-\theta)\left[\kappa_1\left(1-\frac{1}{\psi}\right)\frac{\varphi_e}{1-\kappa_1\rho}\right], \quad \text{and } \lambda_{m,\omega} = (1-\theta)A_2\kappa_1.$

We simulate the dynamics of $(m_{t+1}, \Delta c_{t+1}, \Delta d_{t+1}, x_t, \sigma_t^2)$ using the parameter choices summarized in Panel B of Table IA.I, following exactly the same calibration as in Bansal et al. (2012).

We first simulate the monthly sequences from each model and aggregate them into quarterly observations. Using the quarterly data, we calculate the unconditional risk premia of consumption (Δc) and dividend growth (Δd) as follows:

$$\lambda_g^S = -\frac{\mathbb{E}\left[\text{cov}_t\left(\bar{m}_{t\to t+S},\ g_{t\to t+S}\right)\right]}{S\cdot\sigma(g_{t+1})}, \tag{IA.11}$$

where $\bar{m}_{t+1} = m_{t+1} - \mathbb{E}_t(m_{t+1})$, and we divide the covariance term by $\sigma(g_{t+1})$ because we always

Table IA.I: Parameter Choices in Calibration

Parameter	Variable	Value
Panel A. Campbell and Cochrane (1999)		
Mean consumption growth (%)	g	1.89
Standard deviation of consumption growth (%)	σ	1.50
Log risk-free rate (%)	r_f	0.94
Persistence coefficient	ϕ	0.87
Utility curvature	γ	2.00
Standard deviation of dividend growth (%)	σ_w	11.2
Correlation between Δc and Δd	ρ	0.2
Subjective discount factor	δ	0.89
Steady-state surplus consumption ratio	$\bar{S}$	0.057
Maximum surplus consumption ratio	S_{max}	0.094
Panel B. Bansal, Kiku, and Yaron (2012)		
Subjective discount factor	δ	0.9989
Risk-aversion parameter	γ	10
IES parameter	ψ	1.5
Unconditional mean of consumption growth	μ	0.0015
Persistence coefficient in x_t	ρ	0.975
Persistence coefficient in σ_t^2	ν_1	0.999
Unconditional volatility	σ	0.0072
x_{t+1}'s loading on $\sigma_t e_{t+1}$	φ_e	0.038
σ_t^2's loading on w_{t+1}	σ_w	0.0000028
Unconditional mean of dividend growth	μ_d	0.0015
Δd_{t+1}'s loading on x_t	ϕ_d	2.5
Δd_{t+1}'s loading on $\sigma_t \eta_{t+1}$	π	2.6
Δd_{t+1}'s loading on $\sigma_t u_{t+1}$	φ_d	5.96

The table presents the parameter values used in calibrating the term structure of risk premia in canonical macro-finance models. Panel A shows the parameter choices used in Campbell and Cochrane (1999). Panel B displays the parameter values used in Bansal et al. (2012).

normalize the single-period variable to have unit volatility in the empirical analysis.

Using equation (IA.11), we can obtain closed-form solutions for the term structure of consumption and dividend risk premia in the external habit model,

$$\lambda^S_{\Delta c} = \gamma\sigma\big[1 + \mathbb{E}[\lambda(s_t)]\big], \quad \lambda^S_{\Delta d} = \rho\gamma\sigma\big[1 + \mathbb{E}[\lambda(s_t)]\big].^2 \tag{IA.12}$$

Therefore, the habit model implies flat term structures of risk premia for consumption and dividend growth. We obtain a long sequence of $\lambda(s_t)$, numerically approximate $\mathbb{E}[\lambda(s_t)]$, and estimate $\lambda^S_{\Delta c}$ and $\lambda^S_{\Delta d}$. In contrast, we do not have simple closed-form solutions in the long-run

[2]These are monthly risk premia. The quarterly risk premia equal these monthly numbers multiplied by $\sqrt{3}$ due to the normalization.

risk model; hence, we numerically estimate the risk premia through simulations.

In our empirical analysis, we do not consider single-period dividend growth due to the strong seasonality detected in the data. Instead, we calculate the sum of the lagged 12 monthly dividends, denoted by $D_t^{(12m)}$, and calculate its growth rate as $\Delta d_t^{(12m)} = \log(D_t^{(12m)}/D_{t-1}^{(12m)})$. To make our calibration exercise as close as to the empirical analysis as possible, we estimate the risk premia of $\Delta d_t^{(12m)}$ in the habit and long-run risk models.

Why do we use $\bar{m}_{t+1} = m_{t+1} - \mathbb{E}_t(m_{t+1})$ rather than m_{t+1} in equation (IA.11)? Intuitively, $\mathbb{E}_t(m_{t+1})$ captures the information in the risk-free rate, which is removed because we study the risk premia/average excess returns. In our empirical analysis, we always normalize the log SDF such that its unconditional and conditional means are constant. Using $\bar{m}_{t+1}$ to define risk premia is consistent with our empirical strategy. We now formally show that we can ignore $\mathbb{E}_t(m_{t+1})$ in equation (IA.11) assuming log normality.

Suppose that $R_{t\to t+S} = \prod_{\tau=1}^{S} R_{t+\tau-1\to t+\tau}$ denotes the cumulative gross stock return, $R_{f,t\to t+S} = \prod_{\tau=1}^{S} R_{f,t+\tau-1\to t+\tau}$ denotes the gross risk-free rate, and $M_{t,t+S} = \prod_{\tau=1}^{S} M_{t+\tau-1\to t+\tau}$ is the multi-period SDF that prices the multi-period stock return $R_{t\to t+S}$,

$$\mathbb{E}_t\big[M_{t,t+S}R_{t\to t+S}\big] = \mathbb{E}_t\left[\prod_{\tau=1}^{S} M_{t+\tau-1\to t+\tau}R_{t+\tau-1\to t+\tau}\right] = 1.$$

Define $\tilde{M}_{t+\tau-1\to t+\tau} = \frac{M_{t+\tau-1\to t+\tau}}{\mathbb{E}_{t+\tau-1}[M_{t+\tau-1\to t+\tau}]} = M_{t+\tau-1\to t+\tau}\cdot R_{f,t+\tau-1\to t+\tau}$, where $\mathbb{E}\big[\tilde{M}_{t+\tau-1\to t+\tau}\big] = \mathbb{E}_{t+\tau-1}\big[\tilde{M}_{t+\tau-1\to t+\tau}\big] = 1$. We can rewrite the fundamental asset pricing equation as,

$$\mathbb{E}_t\left[\prod_{\tau=1}^{S} M_{t+\tau-1\to t+\tau}R_{t+\tau-1\to t+\tau}\right] = \mathbb{E}_t\left[\prod_{\tau=1}^{S} \tilde{M}_{t+\tau-1\to t+\tau}\frac{R_{t+\tau-1\to t+\tau}}{R_{f,t+\tau-1\to t+\tau}}\right] = 1, \text{ which implies}$$

$$\mathbb{E}_t\left[\prod_{\tau=1}^{S} \frac{R_{t+\tau-1\to t+\tau}}{R_{f,t+\tau-1\to t+\tau}}\right] - 1 = -\mathrm{cov}_t\left[\prod_{\tau=1}^{S} \tilde{M}_{t+\tau-1\to t+\tau}, \prod_{\tau=1}^{S} \frac{R_{t+\tau-1\to t+\tau}}{R_{f,t+\tau-1\to t+\tau}}\right], \tag{IA.13}$$

where the left side is the multi-horizon excess stock return, and the right side is the covariance between the demeaned cumulative SDF and the excess return.

We now assume that asset returns, macro variables, and the SDF follow log-normal distributions and represent all variables in log units, as follows:

$$\prod_{\tau=1}^{S} \tilde{M}_{t+\tau-1\to t+\tau} = \exp\left\{\sum_{\tau=1}^{S} \tilde{m}_{t+\tau-1\to t+\tau}\right\} = \exp\{\tilde{m}_{t\to t+S}\}, \quad \prod_{\tau=1}^{S} \frac{R_{t+\tau-1\to t+\tau}}{R_{f,t+\tau-1\to t+\tau}} = \exp\{\tilde{r}^e_{t\to t+S}\},$$

$$\mathbb{E}_t\left[\prod_{\tau=1}^{S} \frac{R_{t+\tau-1\to t+\tau}}{R_{f,t+\tau-1\to t+\tau}}\right] - 1 = \mathbb{E}_t\big[\exp\{\tilde{r}^e_{t\to t+S}\}\big] - 1 = \exp\left\{\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big) + \frac{1}{2}\mathrm{var}_t\big(\tilde{r}^e_{t\to t+S}\big)\right\} - 1, \text{ and}$$

$$
\begin{aligned}
&\operatorname{cov}_t\left[\prod_{\tau=1}^{S}\tilde{M}_{t+\tau-1\to t+\tau},\prod_{\tau=1}^{S}\frac{R_{t+\tau-1\to t+\tau}}{R_{f,t+\tau-1\to t+\tau}}\right]=\operatorname{cov}_t\Big[\exp\{\tilde{m}_{t\to t+S}\},\exp\{\tilde{r}^e_{t\to t+S}\}\Big]\\
&=\mathbb{E}_t\big[\exp\{\tilde{r}^e_{t\to t+S}+\tilde{m}_{t\to t+S}\}\big]-\mathbb{E}_t\big[\exp\{\tilde{r}^e_{t\to t+S}\}\big]\cdot\mathbb{E}_t\big[\exp\{\tilde{m}_{t\to t+S}\}\big]\\
&=\exp\left\{\mathbb{E}_t\big(\tilde{m}_{t\to t+S}\big)+\frac{1}{2}\operatorname{var}_t\big(\tilde{m}_{t\to t+S}\big)+\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big)+\frac{1}{2}\operatorname{var}_t\big(\tilde{r}^e_{t\to t+S}\big)+\operatorname{cov}_t\big(\tilde{m}_{t\to t+S},\tilde{r}^e_{t\to t+S}\big)\right\}\\
&\quad-\exp\left\{\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big)+\frac{1}{2}\operatorname{var}_t\big(\tilde{r}^e_{t\to t+S}\big)\big)\right\}\\
&=\exp\left\{\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big)+\frac{1}{2}\operatorname{var}_t\big(\tilde{r}^e_{t\to t+S}\big)+\operatorname{cov}_t\big(\tilde{m}_{t\to t+S},\tilde{r}^e_{t\to t+S}\big)\right\}-\exp\left\{\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big)+\frac{1}{2}\operatorname{var}_t\big(\tilde{r}^e_{t\to t+S}\big)\big)\right\}.
\end{aligned}
$$

In the derivation above, we use the fact that $\exp\{\mathbb{E}_t\big(\tilde{m}_{t\to t+S}\big)+\frac{1}{2}\operatorname{var}_t\big(\tilde{m}_{t\to t+S}\big)\}=\mathbb{E}_t\big(\tilde{M}_{t\to t+S}\big)=1$. We remove the common component, $\exp\{\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big)+\frac{1}{2}\operatorname{var}_t\big(\tilde{r}^e_{t\to t+S}\big)\}$, from both the left and right sides of equation (IA.13). We have the following equation:

$$
\exp\big\{-\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big)-\frac{1}{2}\operatorname{var}_t\big(\tilde{r}^e_{t\to t+S}\big)\big\}-1=\exp\left\{\operatorname{cov}_t\big(\tilde{m}_{t\to t+S},\tilde{r}^e_{t\to t+S}\big)\right\}-1,\ \text{which implies}
$$

$$
\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big)+\frac{1}{2}\operatorname{var}_t\big(\tilde{r}^e_{t\to t+S}\big)=-\operatorname{cov}_t\big(\tilde{m}_{t\to t+S},\tilde{r}^e_{t\to t+S}\big).
$$

Therefore, $-\operatorname{cov}_t\big(\tilde{m}_{t\to t+S},\tilde{r}^e_{t\to t+S}\big)$ properly quantifies the risk premia of the log multi-horizon excess return, conditional on the assumption of log-normality as in our paper and also in many macro-finance models.

We can express $\tilde{m}_{t+1}$ as follows:

$$
\tilde{m}_{t+1}=m_{t+1}-\log(\mathbb{E}_t[M_{t+1}])=m_{t+1}-\mathbb{E}_t[m_{t+1}]-\frac{1}{2}\operatorname{var}_t[m_{t+1}]=\bar{m}_{t+1}-\frac{1}{2}\operatorname{var}_t[m_{t+1}].
$$

It is easy to show that $\operatorname{var}_t[m_{t+1}]$ does not correlate with consumption or dividend growth in the habit and long-run risk formulations that we consider. Finally, dividing the multi-period risk premia, $-\operatorname{cov}_t\big(\tilde{m}_{t\to t+S},\tilde{r}^e_{t\to t+S}\big)$, by the number of periods S, and normalizing by the volatility of the single-period variable, leads to the definition in equation (IA.11).

IA.5 Connection to Dividend Strips

Let D_t denote the dividend payment at time t, and $P_{s,t}$ denotes the time-t price of the dividend strip delivering D_{t+s} at time $t+s$. The fundamental asset pricing equation implies that

$$
P_{s,t}=\mathbb{E}_t[M_{t,t+s}\cdot D_{t+s}], \tag{IA.14}
$$

where $M_{t,t+s}$ is the multi-period SDF between time t and $t+s$. Accordingly, the gross risk-free rate with time-to-maturity s is $R_{f,t,t+s} = 1/\mathbb{E}_t[M_{t,t+s}]$.

The s-period return on the dividend strip with time-to-maturity s can be expressed as

$$R_{t,t+s} = \frac{D_{t+s}}{P_{s,t}} = \frac{D_t}{P_{s,t}} \frac{D_{t+s}}{D_t}, \tag{IA.15}$$

which implies the following per-period log strip return:

$$r_{t+s} = \frac{1}{s}\log(R_{t,t+s}) = \frac{1}{s}\log\left(\frac{D_t}{P_{s,t}}\right) + \frac{1}{s}\log\left(\frac{D_{t+s}}{D_t}\right) = e_{s,t} + g_{d,t,t+s}, \tag{IA.16}$$

where $e_{s,t} = \frac{1}{s}\log\left(\frac{D_t}{P_{s,t}}\right)$ is the spot equity yield for maturity s, and $g_{d,t,t+s} = \frac{1}{s}\log\left(\frac{D_{t+s}}{D_t}\right)$ is the per-period log growth rate of dividend payments. Equation (IA.16) also implies that the conditional variance of r_{t+s} is identical to that of $g_{d,t,t+s}$, that is, $\text{var}_t(r_{t+s}) = \text{var}_t(g_{d,t,t+s})$.

Assuming the joint log normality of dividend growth and the SDF, we can rewrite $\frac{P_{s,t}}{D_t}$ and $e_{s,t}$ as follows:

$$\frac{P_{s,t}}{D_t} = \mathbb{E}_t\left[e^{m_{t,t+s}+\Delta d_{t,t+s}}\right] = \exp\left\{\mathbb{E}_t[m_{t,t+s}]+\frac{1}{2}\text{var}_t(m_{t,t+s})+\mathbb{E}_t[\Delta d_{t,t+s}]+\frac{1}{2}\text{var}_t(\Delta d_{t,t+s})+\text{cov}_t(m_{t,t+s},\Delta d_{t,t+s})\right\},$$

which implies that the spot equity yield, $e_{s,t}$, can be expressed as

$$e_{s,t} = -\frac{1}{s}\left[\mathbb{E}_t[m_{t,t+s}] + \frac{1}{2}\text{var}_t(m_{t,t+s})\right] - \frac{1}{s}\mathbb{E}_t[\Delta d_{t,t+s}] - \frac{1}{2s}\text{var}_t(\Delta d_{t,t+s}) - \frac{1}{s}\text{cov}_t(m_{t,t+s},\Delta d_{t,t+s}).$$

This further implies that the expected per-period strip return is

$$\mathbb{E}_t[r_{t+s}] = e_{s,t} + \mathbb{E}_t[g_{d,t,t+s}] = \frac{1}{s}r_{f,t,t+s} - \frac{1}{2s}\text{var}_t(\Delta d_{t,t+s}) - \frac{1}{s}\text{cov}_t(m_{t,t+s},\Delta d_{t,t+s}). \tag{IA.17}$$

Note that $\text{var}_t(\Delta d_{t,t+s}) = s^2\text{var}_t(r_{t+s})$ and $-\frac{1}{s}\text{cov}_t(m_{t,t+s},\Delta d_{t,t+s})$ is the risk premium of the dividend growth defined in equation (21); hence, we can express the expected per-period strip return, after accounting for the Jensen's correction term, as follows:

$$\mathbb{E}_t[r_{t+s}] + \frac{s}{2}\text{var}_t(r_{t+s}) = \frac{1}{s}r_{f,t,t+s} + \lambda^s_{dt}. \tag{IA.18}$$

Note that under the joint log normality assumption, $\mathbb{E}_t[r_{t+s}] + \frac{s}{2}\text{var}_t(r_{t+s}) = \frac{1}{s}\log\mathbb{E}_t[R_{t,t+s}]$. Therefore, λ^s_{dt}, which equals $\frac{1}{s}\log\mathbb{E}_t[R_{t,t+s}] - \frac{1}{s}r_{f,t,t+s}$, can be interpreted as the per-period risk premium on the hold-to-maturity dividend strips.

Using equation (IA.17), we can derive the forward equity yield, as follows:

$$e^f_{s,t} = e_{s,t} - \frac{1}{s} r_{f,t,t+s} = \lambda^s_{dt} - \mathbb{E}_t[g_{d,t,t+s}] - \frac{1}{2s}\mathrm{var}_t(\Delta d_{t,t+s}), \tag{IA.19}$$

where the last term, $\frac{1}{2s}\mathrm{var}_t(\Delta d_{t,t+s})$ is negligible in the empirical data.

IA.6 Additional Tables

Table IA.II: List of Factors

Number and description of factors:	Sample	Source
AEM intermediary factor (Adrian et al. (2014))	Q1 1968 – Q3 2017	Tyler Muir's Website
Capital share growth (Lettau et al. (2019))	Q3 1963 – Q4 2013	Website of Journal of Finance
Industrial production growth (log change in real per capita)	Q3 1963 – Q4 2019	Federal Reserve Bank of St. Louis
GDP growth (log change in real per capita)	Q3 1963 – Q4 2019	BEA Table 7.1
Durable consumption growth (log change in real per capita)	Q3 1963 – Q4 2019	BEA Table 7.1
Nondurable consumption growth (log change in real per capita)	Q3 1963 – Q4 2019	BEA Table 7.1
Service consumption growth (log change in real per capita)	Q3 1963 – Q4 2019	BEA Table 7.1
Labor income growth (defined in Lettau and Ludvigson (2001))	Q3 1963 – Q3 2019	Martin Lettau's website
Macro PCs 1–5 (FRED-QD, McCracken and Ng (2020))	Q3 1963 – Q4 2019	Michael W. McCracken's website
Oil price (log) change, Spot Crude Oil Price: WTISPLC	Jan 1982 – Dec 2019	Federal Reserve Bank of St. Louis
TED spread (log) change	Jan 1986 – Dec 2019	Federal Reserve Bank of St. Louis
(Non)traded HKM intermediary factors (He et al. (2017))	Jan 1970 – Dec 2019	Zhiguo He's website
PS nontraded liquidity factor (Pástor and Stambaugh (2003))	Jul 1963 – Dec 2019	Lubos Pastor's website
$\Delta \log(\mathrm{VIX}_t) = \log(\mathrm{VIX}_t) - \log(\mathrm{VIX}_{t-1})$	Jan 1986 – Dec 2019	Federal Reserve Bank of St. Louis
Real dividend (log) growth of the S&P500 index	Q3 1963 – Q4 2019	Robert Shiller's website
Price-earning ratio of the S&P500 index (PE_{t-1})	Q3 1963 – Q4 2019	Robert Shiller's website
Term spread (TS_{t-1}) from FRED-QD/MD	Q3 1963 – Q4 2019	Michael W. McCracken's website
Default spread (DS_{t-1}) from FRED-QD/MD	Q3 1963 – Q4 2019	Michael W. McCracken's website
Value spread (VS_{t-1})	Q3 1963 – Q4 2019	Ken French's website
MKT (market), SMB (size), HML (value), MOM (momentum)	Jul 1963 – Dec 2019	Ken French's website

The table presents a list of factors used in Section 4. For each variable, we show the name, description, sample, and data source. In particular, we download the monthly real dividend payments of the S&P500 index from Robert Shiller's website. To avoid the mechanical seasonality in dividend payments, we first calculate the sum of the lagged 12 monthly dividends, denoted by D_t, and compute its growth rate as $\log(D_t/D_{t-1})$. Term spread is the difference between the 10-year and three-month Treasury yields. Default spread is the difference between the yields of the BAA and AAA corporate bonds. The value spread is constructed following Campbell and Vuolteenaho (2004) and Campbell et al. (2013).

Table IA.III: Testing risk premia of strong factors at monthly frequencies ($T = 600$)

	$S=0$	2	4	6	8	10	12	14	16	18	20	22	24
Panel A: $R_g^2 = 30\%$													
Number of Factors = 5													
10%	0.075	0.112	0.102	0.104	0.103	0.099	0.103	0.103	0.101	0.102	0.099	0.102	0.095
5%	0.024	0.062	0.047	0.046	0.050	0.045	0.045	0.046	0.044	0.043	0.044	0.051	0.049
1%	0.002	0.016	0.019	0.015	0.015	0.014	0.015	0.013	0.013	0.014	0.013	0.013	0.013
Number of Factors = 4													
10%	0.030	0.395	0.442	0.442	0.451	0.448	0.449	0.447	0.442	0.443	0.448	0.445	0.446
5%	0.007	0.293	0.331	0.333	0.338	0.327	0.327	0.332	0.332	0.330	0.331	0.330	0.333
1%	0.001	0.134	0.158	0.156	0.153	0.146	0.146	0.149	0.152	0.150	0.147	0.150	0.147
Number of Factors = 7													
10%	0.070	0.110	0.100	0.102	0.106	0.102	0.105	0.101	0.099	0.098	0.093	0.097	0.094
5%	0.020	0.063	0.051	0.047	0.046	0.044	0.044	0.044	0.046	0.045	0.043	0.050	0.052
1%	0.001	0.016	0.017	0.015	0.015	0.017	0.016	0.014	0.016	0.016	0.016	0.015	0.016
Panel B: $R_g^2 = 20\%$													
Number of Factors = 5													
10%	0.059	0.111	0.107	0.094	0.096	0.094	0.091	0.090	0.094	0.089	0.093	0.092	0.090
5%	0.028	0.061	0.051	0.054	0.049	0.049	0.044	0.047	0.052	0.053	0.056	0.053	0.051
1%	0.004	0.008	0.011	0.009	0.011	0.009	0.008	0.008	0.010	0.009	0.012	0.011	0.011
Number of Factors = 4													
10%	0.026	0.390	0.432	0.421	0.420	0.424	0.417	0.429	0.422	0.427	0.428	0.438	0.435
5%	0.008	0.275	0.312	0.311	0.310	0.308	0.312	0.316	0.312	0.303	0.310	0.309	0.305
1%	0.000	0.111	0.133	0.131	0.130	0.134	0.136	0.138	0.142	0.134	0.136	0.134	0.139
Number of Factors = 7													
10%	0.051	0.126	0.102	0.101	0.097	0.096	0.092	0.092	0.091	0.090	0.091	0.089	0.092
5%	0.026	0.062	0.052	0.053	0.053	0.052	0.051	0.051	0.048	0.052	0.056	0.056	0.056
1%	0.003	0.009	0.011	0.010	0.011	0.012	0.010	0.011	0.010	0.011	0.013	0.011	0.010
Panel C: $R_g^2 = 10\%$													
Number of Factors = 5													
10%	0.042	0.149	0.149	0.136	0.133	0.140	0.133	0.142	0.137	0.136	0.134	0.134	0.139
5%	0.017	0.070	0.076	0.086	0.082	0.083	0.085	0.082	0.082	0.083	0.077	0.079	0.090
1%	0.003	0.007	0.024	0.021	0.025	0.019	0.021	0.021	0.017	0.017	0.019	0.020	0.022
Number of Factors = 4													
10%	0.018	0.313	0.373	0.376	0.382	0.384	0.378	0.386	0.384	0.379	0.375	0.369	0.366
5%	0.004	0.191	0.258	0.255	0.269	0.269	0.264	0.261	0.267	0.269	0.268	0.261	0.261
1%	0.002	0.037	0.110	0.111	0.117	0.119	0.121	0.121	0.117	0.108	0.107	0.108	0.112
Number of Factors = 7													
10%	0.039	0.144	0.155	0.150	0.142	0.136	0.140	0.138	0.143	0.138	0.146	0.136	0.140
5%	0.013	0.075	0.089	0.086	0.088	0.093	0.093	0.089	0.090	0.093	0.087	0.087	0.087
1%	0.002	0.006	0.029	0.026	0.026	0.025	0.025	0.025	0.020	0.022	0.026	0.026	0.027

The table reports the frequency of rejecting the null hypothesis $H_0 : \lambda_g^S = \lambda_g^{S,\star}$ based on the 90%, 95%, and 99% credible intervals of our Bayesian estimates in Proposition 1. λ_g^S is defined in equation (6), and $\lambda_g^{S,\star}$ is λ_g^S's pseudo-true value. We consider strong factors, with $R_g^2 \in \{10\%, 20\%, 30\%\}$. We simulate monthly observations of g_t and $\boldsymbol{r}_t$ by assuming that i) the true number of latent factors is 5, ii) the time series sample size is 600 quarters, and iii) the true $\bar{S} = 16$. We estimate several model configurations with different numbers of factors (4, 5, and 7), and $\bar{S} = 24$. The number of Monte Carlo simulations is 1,000.

Table IA.IV: Testing risk premia of useless factors at quarterly frequencies ($T = 200$)

	$S = 0$	1	2	3	4	5	6	7	8	9	10	11	12
Panel A: $R_g^2 = 30\%$													
Number of Factors = 5													
10%	0.032	0.036	0.043	0.047	0.052	0.051	0.057	0.056	0.060	0.067	0.072	0.073	0.073
5%	0.016	0.015	0.014	0.019	0.022	0.027	0.027	0.029	0.036	0.034	0.039	0.038	0.039
1%	0.003	0.001	0.002	0.005	0.007	0.008	0.008	0.007	0.010	0.012	0.012	0.012	0.012
Number of Factors = 4													
10%	0.027	0.038	0.048	0.049	0.051	0.048	0.047	0.053	0.052	0.052	0.059	0.063	0.065
5%	0.011	0.016	0.021	0.027	0.027	0.026	0.027	0.027	0.031	0.028	0.030	0.030	0.031
1%	0.001	0.000	0.003	0.006	0.005	0.004	0.008	0.004	0.006	0.006	0.008	0.009	0.009
Number of Factors = 7													
10%	0.023	0.029	0.040	0.040	0.046	0.053	0.049	0.054	0.054	0.056	0.060	0.062	0.063
5%	0.008	0.009	0.017	0.022	0.024	0.028	0.027	0.028	0.029	0.031	0.033	0.038	0.034
1%	0.002	0.001	0.002	0.004	0.005	0.006	0.006	0.012	0.009	0.008	0.009	0.009	0.010
Panel B: $R_g^2 = 20\%$													
Number of Factors = 5													
10%	0.015	0.025	0.029	0.036	0.038	0.043	0.047	0.049	0.050	0.054	0.050	0.053	0.053
5%	0.006	0.011	0.013	0.014	0.019	0.021	0.021	0.025	0.023	0.025	0.028	0.028	0.030
1%	0.002	0.003	0.003	0.007	0.006	0.004	0.006	0.005	0.005	0.007	0.005	0.004	0.004
Number of Factors = 4													
10%	0.027	0.028	0.034	0.040	0.043	0.043	0.045	0.047	0.050	0.053	0.053	0.051	0.051
5%	0.012	0.011	0.010	0.018	0.017	0.018	0.017	0.019	0.021	0.025	0.026	0.027	0.027
1%	0.001	0.002	0.002	0.002	0.001	0.003	0.003	0.004	0.004	0.005	0.005	0.006	0.004
Number of Factors = 7													
10%	0.011	0.022	0.023	0.028	0.036	0.039	0.039	0.045	0.050	0.051	0.055	0.052	0.050
5%	0.006	0.008	0.013	0.018	0.018	0.022	0.023	0.023	0.024	0.026	0.023	0.022	0.019
1%	0.002	0.001	0.003	0.004	0.004	0.003	0.004	0.004	0.004	0.005	0.005	0.005	0.006
Panel C: $R_g^2 = 10\%$													
Number of Factors = 5													
10%	0.026	0.022	0.027	0.030	0.033	0.034	0.033	0.033	0.032	0.031	0.033	0.036	0.035
5%	0.010	0.007	0.014	0.014	0.015	0.014	0.016	0.018	0.017	0.017	0.018	0.018	0.016
1%	0.001	0.000	0.000	0.002	0.002	0.001	0.002	0.002	0.003	0.000	0.002	0.003	0.003
Number of Factors = 4													
10%	0.023	0.024	0.027	0.029	0.036	0.033	0.027	0.031	0.026	0.032	0.031	0.031	0.033
5%	0.010	0.010	0.019	0.016	0.019	0.013	0.013	0.014	0.013	0.013	0.015	0.015	0.014
1%	0.002	0.003	0.001	0.003	0.003	0.004	0.004	0.005	0.004	0.003	0.004	0.003	0.003
Number of Factors = 7													
10%	0.025	0.017	0.021	0.022	0.025	0.019	0.022	0.021	0.021	0.019	0.021	0.025	0.022
5%	0.005	0.006	0.009	0.008	0.010	0.008	0.013	0.013	0.011	0.010	0.014	0.014	0.016
1%	0.000	0.000	0.000	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.000	0.000	0.002

The table reports the frequency of rejecting the null hypothesis $H_0 : \lambda_g^S = 0$ based on the 90%, 95%, and 99% credible intervals of our Bayesian estimates in Proposition 1. λ_g^S is defined in equation (6). We consider useless factors with different degrees of persistency; that is, the persistent component in g_t accounts for 10%, 20%, or 30% of time series variations. We simulate quarterly observations of g_t and $\boldsymbol{r}_t$ by assuming that i) the true number of latent factors is 5, ii) the time series sample size is 200 quarters, and iii) g_t is orthogonal to $\boldsymbol{r}_t$. We estimate several model configurations with different numbers of factors (4, 5, and 7), and $\bar{S} = 12$. The number of Monte Carlo simulations is 1,000.

Table IA.V: Testing risk premia of useless factors at monthly frequencies ($T = 600$)

	$S = 0$	2	4	6	8	10	12	14	16	18	20	22	24
Panel A: $R_g^2 = 30\%$													
Number of Factors = 5													
10%	0.019	0.038	0.046	0.051	0.055	0.061	0.066	0.078	0.083	0.085	0.086	0.090	0.089
5%	0.006	0.013	0.017	0.024	0.029	0.033	0.035	0.035	0.046	0.048	0.049	0.053	0.049
1%	0.001	0.001	0.002	0.003	0.006	0.006	0.007	0.007	0.009	0.013	0.013	0.014	0.013
Number of Factors = 4													
10%	0.015	0.032	0.044	0.054	0.059	0.065	0.076	0.082	0.086	0.088	0.093	0.094	0.089
5%	0.009	0.013	0.019	0.024	0.032	0.033	0.039	0.043	0.045	0.051	0.051	0.049	0.050
1%	0.000	0.002	0.003	0.006	0.009	0.011	0.013	0.011	0.010	0.011	0.015	0.013	0.013
Number of Factors = 7													
10%	0.013	0.024	0.029	0.041	0.050	0.053	0.063	0.065	0.075	0.077	0.089	0.088	0.082
5%	0.005	0.014	0.016	0.022	0.025	0.026	0.029	0.034	0.038	0.043	0.043	0.043	0.044
1%	0.000	0.000	0.002	0.004	0.005	0.006	0.007	0.007	0.008	0.011	0.014	0.016	0.017
Panel B: $R_g^2 = 20\%$													
Number of Factors = 5													
10%	0.017	0.035	0.040	0.056	0.060	0.060	0.073	0.072	0.070	0.076	0.073	0.076	0.073
5%	0.003	0.016	0.026	0.027	0.040	0.040	0.035	0.042	0.045	0.045	0.042	0.044	0.042
1%	0.001	0.005	0.007	0.006	0.012	0.010	0.009	0.008	0.011	0.013	0.015	0.014	0.014
Number of Factors = 4													
10%	0.020	0.037	0.052	0.052	0.062	0.068	0.072	0.075	0.076	0.079	0.074	0.081	0.079
5%	0.007	0.013	0.026	0.030	0.034	0.039	0.035	0.039	0.037	0.042	0.040	0.044	0.042
1%	0.001	0.004	0.004	0.005	0.008	0.009	0.008	0.007	0.012	0.010	0.009	0.010	0.010
Number of Factors = 7													
10%	0.009	0.026	0.034	0.036	0.046	0.045	0.053	0.057	0.053	0.054	0.058	0.058	0.063
5%	0.003	0.012	0.018	0.018	0.020	0.023	0.027	0.030	0.036	0.035	0.036	0.041	0.040
1%	0.000	0.003	0.003	0.004	0.006	0.007	0.007	0.008	0.006	0.009	0.010	0.007	0.008
Panel C: $R_g^2 = 10\%$													
Number of Factors = 5													
10%	0.011	0.021	0.024	0.032	0.038	0.040	0.044	0.042	0.043	0.043	0.045	0.046	0.046
5%	0.003	0.005	0.008	0.010	0.013	0.011	0.014	0.021	0.017	0.018	0.018	0.022	0.021
1%	0.000	0.000	0.001	0.001	0.001	0.001	0.003	0.002	0.003	0.002	0.003	0.008	0.006
Number of Factors = 4													
10%	0.018	0.029	0.028	0.035	0.037	0.037	0.041	0.047	0.049	0.045	0.054	0.044	0.040
5%	0.005	0.010	0.006	0.015	0.015	0.013	0.015	0.019	0.022	0.021	0.018	0.019	0.020
1%	0.000	0.001	0.002	0.001	0.002	0.002	0.001	0.001	0.001	0.002	0.004	0.004	0.003
Number of Factors = 7													
10%	0.007	0.010	0.015	0.018	0.023	0.028	0.032	0.032	0.035	0.031	0.034	0.037	0.039
5%	0.001	0.005	0.003	0.007	0.009	0.007	0.010	0.010	0.012	0.009	0.010	0.013	0.013
1%	0.000	0.000	0.001	0.001	0.001	0.002	0.001	0.000	0.002	0.002	0.002	0.002	0.002

The table reports the frequency of rejecting the null hypothesis $H_0 : \lambda_g^S = 0$ based on the 90%, 95%, and 99% credible intervals of our Bayesian estimates in Proposition 1. λ_g^S is defined in equation (6). We consider useless factors with different degrees of persistency; that is, the persistent component in g_t accounts for 10%, 20%, or 30% of time series variations. We simulate quarterly observations of g_t and $\boldsymbol{r}_t$ by assuming that i) the true number of latent factors is 5, ii) the time series sample size is 600 months, and iii) g_t is orthogonal to $\boldsymbol{r}_t$. We estimate several model configurations with different numbers of factors (4, 5, and 7), and $\bar{S} = 24$. The number of Monte Carlo simulations is 1,000.

Table IA.VI: Bayesian estimates of R_g^2 and $\text{corr}(\hat{f}_t, f_t)$ for strong and useless factors

Number of factors:	$K=4$			$K=5$			$K=7$		
True $R_g^2 =$	10%	20%	30%	10%	20%	30%	10%	20%	30%
Panel A. Posterior distributions of R_g^2									
T = 200, strong factors									
median	0.122	0.187	0.259	0.142	0.224	0.311	0.156	0.235	0.320
5th	0.065	0.109	0.155	0.081	0.142	0.206	0.094	0.152	0.217
95th	0.199	0.288	0.370	0.228	0.325	0.421	0.240	0.337	0.428
T = 600, strong factors									
median	0.104	0.185	0.272	0.112	0.204	0.297	0.116	0.207	0.300
5th	0.064	0.130	0.201	0.073	0.148	0.227	0.074	0.151	0.231
95th	0.149	0.249	0.344	0.158	0.264	0.370	0.163	0.267	0.373
T = 200, useless factors									
median	0.081	0.083	0.084	0.091	0.094	0.097	0.109	0.113	0.117
5th	0.045	0.046	0.045	0.054	0.055	0.054	0.069	0.072	0.073
95th	0.133	0.140	0.145	0.145	0.154	0.167	0.165	0.178	0.195
T = 600, useless factors									
median	0.041	0.042	0.044	0.046	0.047	0.049	0.052	0.055	0.059
5th	0.027	0.026	0.026	0.030	0.031	0.030	0.035	0.037	0.037
95th	0.063	0.071	0.080	0.067	0.077	0.087	0.075	0.086	0.102
Panel B. Posterior distributions of $\text{corr}(\hat{f}_t, f_t)$									
T = 200, strong factors									
median	0.702	0.810	0.842	0.773	0.902	0.937	0.677	0.855	0.910
5th	0.324	0.605	0.736	0.386	0.745	0.860	0.294	0.665	0.809
95th	0.849	0.883	0.895	0.910	0.951	0.967	0.864	0.926	0.951
T = 600, strong factors									
median	0.885	0.916	0.927	0.919	0.960	0.972	0.882	0.944	0.963
5th	0.760	0.873	0.893	0.812	0.925	0.953	0.747	0.901	0.940
95th	0.928	0.944	0.950	0.955	0.974	0.980	0.934	0.964	0.974

The table reports the Bayesian estimates of R_g^2 and $\text{corr}(\hat{f}_t, f_t)$ for strong and useless factors: (1) R_g^2 measures the percentage of g_t's time series variations explained by asset returns' latent factors, and (2) $\text{corr}(\hat{f}_t, f_t)$ quantifies the correlation between the true f_t and its estimate, $\hat{f}_t = \hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{v}}_t$. For useless factors, we report only R_g^2. In each model, we report the median, 5th, and 95th percentiles based on 1,000 simulations. We consider strong and useless factors with different degrees of persistency; that is, the persistent component in g_t accounts for 10%, 20%, or 30% of time series variations. We simulate monthly or quarterly observations of g_t and $\boldsymbol{r}_t$ by assuming that the true number of latent factors is 5. We estimate several model configurations with different numbers of factors ($K \in \{4, 5, 7\}$), and $\bar{S} = 12$ for $T = 200$ ($\bar{S} = 24$ for $T = 600$).

Table IA.VII: Size and power of the Bayesian estimates and Giglio and Xiu (2021)

	Bayesian Estimation						Giglio and Xiu (2021)					
	Five factors			Seven factors			Five factors			Seven factors		
	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
Panel A. Size												
$T = 200$												
10%	0.066	0.030	0.012	0.068	0.026	0.009	0.072	0.033	0.012	0.075	0.029	0.012
20%	0.087	0.042	0.008	0.085	0.045	0.005	0.091	0.048	0.008	0.082	0.046	0.005
30%	0.095	0.058	0.008	0.089	0.053	0.008	0.096	0.059	0.009	0.093	0.055	0.008
$T = 600$												
10%	0.101	0.047	0.009	0.098	0.043	0.010	0.103	0.053	0.007	0.111	0.041	0.007
20%	0.100	0.050	0.011	0.097	0.048	0.008	0.099	0.051	0.009	0.100	0.056	0.008
30%	0.104	0.050	0.016	0.106	0.048	0.016	0.110	0.048	0.013	0.099	0.050	0.012
Panel B. Power												
$T = 200$												
10%	0.278	0.190	0.051	0.267	0.160	0.046	0.286	0.188	0.045	0.288	0.189	0.040
20%	0.403	0.279	0.119	0.387	0.265	0.101	0.397	0.282	0.101	0.396	0.273	0.099
30%	0.484	0.371	0.169	0.466	0.358	0.154	0.478	0.359	0.151	0.480	0.370	0.155
$T = 600$												
10%	0.520	0.410	0.186	0.499	0.391	0.189	0.523	0.414	0.174	0.507	0.391	0.176
20%	0.658	0.545	0.307	0.659	0.530	0.298	0.652	0.540	0.295	0.649	0.530	0.287
30%	0.715	0.598	0.365	0.708	0.583	0.364	0.711	0.590	0.343	0.699	0.581	0.334

Panel A reports the frequency of rejecting the null hypothesis $H_0 : \lambda_g = \lambda_g^\star$ based on the 90%, 95%, and 99% credible intervals given by (1) our Bayesian estimates in Proposition 1 and (2) the frequentist test statistics in Theorem 1 of Giglio and Xiu (2021). $\lambda_g^\star$ is λ_g's pseudo-true value. Differently, Panel B reports the frequency of rejecting the null hypothesis $H_0 : \lambda_g = 0$. We consider strong factors, with $R_g^2 \in \{10\%, 20\%, 30\%\}$. We simulate quarterly ($T = 200$) and monthly ($T = 600$) observations of g_t and $\boldsymbol{r}_t$ by assuming that i) the true number of latent factors is five and ii) g_t correlates with only on the contemporaneous $\tilde{\boldsymbol{v}}_t$ ($\bar{S} = 0$). We estimate several model configurations with different numbers of factors (5, 7). The number of Monte Carlo simulations is 1,000.

Table IA.VIII: Bayesian estimates of R_g^2 and $\text{corr}(\hat{f}_t, f_t)$ for strong and useless factors

	Panel A. R_g^2						**Panel B.** $\text{corr}(\hat{f}_t, f_t)$					
Number of factors:	$K = 5$			$K = 7$			$K = 5$			$K = 7$		
True $R_g^2 =$	10%	20%	30%	10%	20%	30%	10%	20%	30%	10%	20%	30%
Quarterly frequency ($T = 200$)												
median	0.146	0.223	0.310	0.157	0.232	0.318	0.762	0.883	0.921	0.677	0.831	0.888
5th	0.081	0.139	0.208	0.097	0.144	0.219	0.343	0.724	0.846	0.283	0.635	0.795
95th	0.219	0.328	0.424	0.229	0.333	0.428	0.893	0.936	0.953	0.842	0.905	0.931
Monthly frequency ($T = 600$)												
median	0.115	0.208	0.299	0.118	0.210	0.301	0.913	0.953	0.965	0.874	0.935	0.953
5th	0.069	0.148	0.227	0.074	0.151	0.229	0.794	0.916	0.942	0.717	0.888	0.928
95th	0.165	0.270	0.376	0.168	0.272	0.377	0.951	0.969	0.976	0.926	0.957	0.968

The table reports the Bayesian estimates of R_g^2 and $\text{corr}(\hat{f}_t, f_t)$ for strong factors: (1) R_g^2 measures the percentage of g_t's time series variations explained by $\boldsymbol{\epsilon}_{vt}$, and (2) $\text{corr}(\hat{f}_t, f_t)$ quantifies the correlation between the true f_t and its estimate, $\hat{f}_t = \hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{\epsilon}}_{vt}$. In each model, we report the median, 5th, and 95th percentiles based on 1,000 simulations. We consider several degrees of persistency; that is, the persistent component in g_t accounts for 10%, 20%, or 30% of time series variations. We simulate monthly or quarterly observations of g_t and $\boldsymbol{r}_t$ by assuming that the true number of latent factors is 5. We estimate several model configurations with different numbers of factors ($K \in \{5, 7\}$), and $\bar{S} = 12$ for $T = 200$ ($\bar{S} = 24$ for $T = 600$).

Table IA.IX: Testing unconditional risk premia of strong factors at quarterly frequencies ($T = 200$) when factors command time-varying risk premia in simulations

	$S = 0$	1	2	3	4	5	6	7	8	9	10	11	12
							Panel A: $R_g^2 = 30\%$						
							Number of Factors = 5						
10%	0.088	0.099	0.102	0.111	0.115	0.115	0.111	0.114	0.113	0.115	0.111	0.113	0.112
5%	0.041	0.054	0.058	0.059	0.060	0.060	0.061	0.059	0.059	0.059	0.058	0.058	0.056
1%	0.010	0.013	0.017	0.015	0.015	0.016	0.016	0.017	0.019	0.017	0.017	0.018	0.019
							Number of Factors = 7						
10%	0.092	0.103	0.108	0.112	0.114	0.111	0.111	0.107	0.106	0.107	0.105	0.102	0.095
5%	0.048	0.050	0.056	0.056	0.053	0.055	0.057	0.055	0.057	0.053	0.054	0.054	0.055
1%	0.010	0.011	0.012	0.015	0.015	0.014	0.014	0.015	0.015	0.013	0.014	0.012	0.014
							Panel B: $R_g^2 = 20\%$						
							Number of Factors = 5						
10%	0.103	0.104	0.103	0.097	0.096	0.100	0.103	0.099	0.103	0.108	0.110	0.114	0.103
5%	0.054	0.048	0.052	0.047	0.048	0.049	0.045	0.049	0.050	0.053	0.050	0.055	0.052
1%	0.008	0.013	0.014	0.011	0.008	0.008	0.011	0.013	0.013	0.012	0.013	0.012	0.013
							Number of Factors = 7						
10%	0.091	0.086	0.086	0.087	0.083	0.082	0.088	0.092	0.091	0.093	0.096	0.097	0.095
5%	0.048	0.045	0.046	0.048	0.045	0.044	0.044	0.043	0.047	0.050	0.054	0.051	0.051
1%	0.005	0.009	0.011	0.010	0.009	0.009	0.012	0.013	0.011	0.012	0.012	0.013	0.012
							Panel C: $R_g^2 = 10\%$						
							Number of Factors = 5						
10%	0.084	0.118	0.121	0.127	0.127	0.132	0.124	0.133	0.132	0.127	0.122	0.124	0.125
5%	0.034	0.064	0.067	0.063	0.070	0.071	0.065	0.067	0.069	0.064	0.068	0.065	0.067
1%	0.007	0.015	0.016	0.015	0.018	0.015	0.012	0.014	0.014	0.014	0.012	0.012	0.012
							Number of Factors = 7						
10%	0.072	0.121	0.113	0.123	0.129	0.127	0.136	0.134	0.137	0.129	0.127	0.126	0.122
5%	0.027	0.066	0.069	0.069	0.072	0.072	0.071	0.064	0.069	0.071	0.069	0.066	0.065
1%	0.004	0.015	0.015	0.012	0.015	0.017	0.016	0.012	0.011	0.011	0.012	0.010	0.011

The table focuses on unconditional risk premia $\lambda_g^S = \sum_{\tau=0}^{S} \sum_{s=0}^{\tau} \frac{\rho_s \boldsymbol{\eta}_g^\top \boldsymbol{\lambda}_v}{1+S}$ and reports the frequency of rejecting the null hypothesis $H_0 : \lambda_g^S = \lambda_g^{S,\star}$ based on the 90%, 95%, and 99% credible intervals of our Bayesian estimates in Proposition A2. $\lambda_g^{S,\star}$ is λ_g^S's pseudo-true value. We consider strong factors, with $R_g^2 \in \{10\%, 20\%, 30\%\}$. We simulate quarterly observations of g_t and $\boldsymbol{r}_t$ by assuming that i) the true number of latent factors is 5, ii) the time series sample size is 200 quarters, iii) the true $\bar{S} = 8$, and iv) the five latent factors follow a VAR(1) process. We estimate several model configurations with different numbers of factors (5, 7), and $\bar{S} = 12$. The number of Monte Carlo simulations is 1,000.

Table IA.X: Testing unconditional risk premia of strong factors at monthly frequencies ($T = 600$) when factors command time-varying risk premia in simulations

	$S=0$	2	4	6	8	10	12	14	16	18	20	22	24
Panel A: $R_g^2 = 30\%$													
Number of Factors = 5													
10%	0.080	0.107	0.127	0.117	0.121	0.121	0.121	0.117	0.119	0.118	0.119	0.115	0.116
5%	0.038	0.054	0.068	0.062	0.063	0.062	0.062	0.065	0.056	0.058	0.056	0.056	0.059
1%	0.005	0.012	0.012	0.016	0.016	0.017	0.017	0.017	0.015	0.016	0.016	0.017	0.017
Number of Factors = 7													
10%	0.077	0.110	0.123	0.115	0.117	0.115	0.120	0.113	0.118	0.115	0.121	0.120	0.114
5%	0.038	0.054	0.062	0.065	0.062	0.058	0.056	0.058	0.053	0.053	0.054	0.052	0.049
1%	0.006	0.006	0.011	0.012	0.013	0.013	0.014	0.012	0.014	0.015	0.013	0.014	0.016
Panel B: $R_g^2 = 20\%$													
Number of Factors = 5													
10%	0.075	0.092	0.107	0.107	0.106	0.109	0.112	0.110	0.111	0.109	0.117	0.112	0.110
5%	0.035	0.050	0.059	0.060	0.062	0.063	0.059	0.065	0.063	0.059	0.058	0.059	0.055
1%	0.004	0.011	0.013	0.011	0.011	0.011	0.012	0.012	0.013	0.013	0.012	0.013	0.013
Number of Factors = 7													
10%	0.069	0.099	0.100	0.105	0.105	0.103	0.103	0.107	0.104	0.102	0.105	0.105	0.109
5%	0.027	0.050	0.054	0.054	0.059	0.057	0.058	0.061	0.060	0.056	0.057	0.057	0.053
1%	0.003	0.008	0.012	0.009	0.013	0.011	0.013	0.013	0.014	0.013	0.011	0.013	0.015
Panel C: $R_g^2 = 10\%$													
Number of Factors = 5													
10%	0.045	0.104	0.102	0.111	0.117	0.116	0.117	0.116	0.116	0.119	0.118	0.110	0.114
5%	0.020	0.049	0.051	0.056	0.052	0.069	0.072	0.068	0.066	0.064	0.068	0.065	0.062
1%	0.004	0.003	0.018	0.014	0.012	0.012	0.017	0.017	0.016	0.017	0.014	0.013	0.011
Number of Factors = 7													
10%	0.044	0.100	0.104	0.106	0.116	0.114	0.120	0.116	0.115	0.123	0.115	0.117	0.113
5%	0.022	0.057	0.052	0.061	0.057	0.060	0.063	0.061	0.060	0.060	0.059	0.064	0.061
1%	0.004	0.004	0.016	0.012	0.011	0.011	0.010	0.011	0.012	0.013	0.013	0.012	0.013

The table focuses on unconditional risk premia $\lambda_g^S = \sum_{\tau=0}^{S} \sum_{s=0}^{\tau} \frac{\rho_s \boldsymbol{\eta}_g^\top \boldsymbol{\lambda}_v}{1+S}$ and reports the frequency of rejecting the null hypothesis $H_0 : \lambda_g^S = \lambda_g^{S,\star}$ based on the 90%, 95%, and 99% credible intervals of our Bayesian estimates in Proposition A2. $\lambda_g^{S,\star}$ is λ_g^S's pseudo-true value. We consider strong factors, with $R_g^2 \in \{10\%, 20\%, 30\%\}$. We simulate monthly observations of g_t and $\boldsymbol{r}_t$ by assuming that i) the true number of latent factors is 5, ii) the time series sample size is 600 quarters, and iii) the true $\bar{S} = 16$, and iv) the five latent factors follow a VAR(1) process. We estimate several model configurations with different numbers of factors (5, 7), and $\bar{S} = 24$. The number of Monte Carlo simulations is 1,000.

Table IA.XI: Factors' risk premia: Six- and seven-factor models

	Panel A. Quarterly variables, $\bar{S} = 12$ quarters							
$S =$	0	2	4	6	8	10	12	R_g^2
	Number of factors = 6							
AEM intermediary	0.103***	0.115***	0.130**	0.124**	0.111	0.092	0.083	15.2%
Capital share growth	0.010	0.010	0.006	0.001	-0.001	-0.002	-0.005	11.1%
GDP growth	0.021	0.065*	0.100*	0.123*	0.136*	0.146*	0.152*	24.1%
IP growth	0.001	0.053*	0.090*	0.111*	0.123*	0.128*	0.130*	40.2%
Durable consumption growth	-0.003	0.075**	0.112**	0.129**	0.136**	0.143**	0.150**	18.4%
Nondurable consumption growth	0.031**	0.082**	0.109**	0.138**	0.160**	0.175**	0.189**	23.3%
Service consumption growth	0.034	0.053	0.074	0.085	0.091	0.100	0.110	11.8%
Nondurable + service	0.041*	0.086*	0.124*	0.153*	0.176*	0.194*	0.212*	18.6%
Labor income growth	0.011	0.007	0.006	0.007	0.008	0.009	0.013	10.5%
Dividend growth of S&P500	0.000	0.020	0.060*	0.106**	0.153**	0.195**	0.232**	44.7%
Macro PC1 (FRED-QD)	0.013	0.076**	0.137**	0.184**	0.220**	0.251**	0.273**	48.7%
Macro PC2 (FRED-QD)	0.063	0.100	0.104	0.092	0.078	0.066	0.053	41.0%
Macro PC3 (FRED-QD)	0.001	-0.001	-0.003	-0.006	-0.012	-0.022	-0.034	17.8%
Macro PC4 (FRED-QD)	-0.132***	-0.150***	-0.200***	-0.255***	-0.305***	-0.353***	-0.397***	48.0%
Macro PC5 (FRED-QD)	0.084**	0.095*	0.079	0.057	0.038	0.024	0.013	30.5%
	Number of factors = 7							
AEM intermediary	0.114***	0.115***	0.123**	0.113**	0.099*	0.079	0.067	16.5%
Capital share growth	0.006	0.006	0.004	0.001	0.000	-0.002	-0.003	10.3%
GDP growth	0.018	0.060*	0.094*	0.116*	0.127*	0.136*	0.142*	24.5%
IP growth	0.002	0.051*	0.086*	0.106*	0.118*	0.124*	0.126*	41.1%
Durable consumption growth	-0.002	0.070**	0.105***	0.123***	0.131***	0.140***	0.146***	18.7%
Nondurable consumption growth	0.027**	0.075**	0.102**	0.130**	0.150**	0.165**	0.179**	24.8%
Service consumption growth	0.027	0.046	0.065	0.075	0.080	0.089	0.099	11.8%
Nondurable + service	0.039*	0.083*	0.120*	0.150*	0.173*	0.191*	0.209**	18.9%
Labor income growth	0.007	0.004	0.004	0.004	0.005	0.007	0.009	11.2%
Dividend growth S&P500	-0.001	0.014	0.051*	0.095**	0.139**	0.179**	0.215**	46.4%
Macro PC1 (FRED-QD)	0.013	0.071**	0.129**	0.176**	0.211**	0.240**	0.263**	48.8%
Macro PC2 (FRED-QD)	0.061	0.095	0.098	0.086	0.074	0.062	0.050	41.4%
Macro PC3 (FRED-QD)	0.001	-0.002	-0.005	-0.008	-0.015	-0.027	-0.041	18.0%
Macro PC4 (FRED-QD)	-0.134***	-0.153***	-0.204***	-0.262***	-0.313***	-0.361***	-0.402***	48.1%
Macro PC5 (FRED-QD)	0.044	0.063	0.061	0.056	0.051	0.047	0.044	36.5%
	Panel B. Monthly variables, $\bar{S} = 24$ months							
$S =$	0	4	8	12	16	20	24	R_g^2
	Number of factors = 6							
Oil price change	-0.018	-0.045*	-0.056*	-0.059*	-0.064*	-0.067*	-0.068*	11.6%
TED spread change	-0.005	-0.007	-0.005	-0.005	-0.005	-0.004	-0.004	12.9%
Nontraded HKM intermediary	0.102***	0.105***	0.102***	0.097***	0.096***	0.094***	0.093***	61.4%
Traded HKM intermediary	0.117***	0.119***	0.114***	0.108***	0.105***	0.102***	0.101***	71.5%
Nontraded liquidity	0.044**	0.065**	0.077**	0.087**	0.096**	0.104**	0.110**	15.4%
$\Delta\log(\text{VIX})$	-0.126***	-0.077***	-0.061***	-0.047***	-0.040***	-0.035***	-0.030***	51.7%
	Number of factors = 7							
Oil price change	-0.016	-0.041*	-0.052*	-0.055*	-0.058*	-0.061*	-0.063*	11.4%
TED spread change	-0.006	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	18.6%
Nontraded HKM intermediary	0.100***	0.103***	0.099***	0.096***	0.094***	0.092***	0.091***	61.5%
Traded HKM intermediary	0.116***	0.118***	0.111***	0.105***	0.102***	0.099***	0.097***	71.9%
PS liquidity	0.043**	0.063**	0.074**	0.083**	0.092**	0.099**	0.105**	15.4%
$\Delta\log(\text{VIX})$	-0.129***	-0.078***	-0.061***	-0.048***	-0.040***	-0.035***	-0.030**	51.6%

The table repeats the same analysis in Table 1 of the main text. However, unlike Table 1, we consider six- and seven-factor models for asset returns in this table.

Table IA.XII: Which principal components of returns drive the common component $\hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{v}}_t$?

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	Total R^2
	Panel A. Quarterly variables, $\bar{S} = 12$ quarters							
AEM intermediary	0.09	0.21	0.42	0.06	0.06	0.07	0.05	0.95
Capital share growth	0.21	0.13	0.13	0.08	0.07	0.28	0.05	0.95
GDP growth	0.66	0.02	0.02	0.07	0.15	0.03	0.03	0.99
IP growth	0.69	0.01	0.01	0.04	0.17	0.04	0.03	0.99
Durable consumption growth	0.38	0.07	0.16	0.18	0.09	0.07	0.03	0.98
Nondurable consumption growth	0.67	0.02	0.06	0.02	0.12	0.04	0.06	0.98
Service consumption growth	0.20	0.07	0.07	0.34	0.10	0.09	0.08	0.94
Nondurable + service	0.53	0.04	0.05	0.15	0.10	0.05	0.05	0.97
Labor income growth	0.05	0.06	0.07	0.13	0.07	0.20	0.28	0.87
Dividend growth of SP500	0.56	0.04	0.07	0.12	0.14	0.01	0.04	0.97
Macro PC1 (FRED-QD)	0.76	0.01	0.01	0.01	0.17	0.01	0.02	0.99
Macro PC2 (FRED-QD)	0.78	0.03	0.01	0.03	0.06	0.07	0.02	1.00
Macro PC3 (FRED-QD)	0.14	0.05	0.07	0.23	0.11	0.28	0.06	0.94
Macro PC4 (FRED-QD)	0.77	0.07	0.07	0.01	0.06	0.01	0.01	0.99
Macro PC5 (FRED-QD)	0.40	0.05	0.05	0.02	0.02	0.13	0.22	0.89
	Panel B. Monthly variables, $\bar{S} = 24$ months							
Oil price change	0.10	0.06	0.16	0.23	0.05	0.09	0.11	0.80
TED spread change	0.11	0.04	0.03	0.13	0.09	0.01	0.14	0.55
Nontraded HKM intermediary	0.77	0.14	0.00	0.06	0.02	0.01	0.00	1.00
Traded HKM intermediary	0.79	0.14	0.00	0.04	0.02	0.01	0.00	1.00
PS liquidity	0.82	0.03	0.02	0.03	0.01	0.05	0.00	0.96
$\Delta \log(\text{VIX})$	0.87	0.06	0.02	0.01	0.00	0.00	0.00	0.96

The table reports the posterior means of the squared correlation between the common component estimates, $\hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{v}}_t$, and the first seven principal components of asset returns. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. In the last column, we also report the sum of the first seven columns and denote it as the total R^2. All variables are standardized to have unit variances. We consider a seven-factor model for asset returns. Panel A tabulates the estimates of quarterly factors, using a lag of 12 quarters in g_t's equations. Panel B tabulates the estimates of monthly factors, using a lag of 24 months in estimation. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

Table IA.XIII: Are MA components of macro factors similar? (six- and seven-factor models)

	GDP growth	IP growth	Durable	Nondurable	Service	Dividend	Macro PC1	Macro PC2	Macro PC4
	Panel A. Number of factors = 6								
GDP growth	1.00	0.89	0.72	0.69	0.48	0.36	0.89	0.47	-0.43
IP growth	0.89	1.00	0.74	0.72	0.38	0.35	0.84	0.42	-0.23
Durable	0.72	0.74	1.00	0.69	0.36	0.31	0.64	0.38	-0.22
Nondurable	0.69	0.72	0.69	1.00	0.45	0.45	0.71	0.41	-0.52
Service	0.48	0.38	0.36	0.45	1.00	0.14	0.48	0.15	-0.33
Dividend growth	0.36	0.35	0.31	0.45	0.14	1.00	0.63	-0.22	-0.59
Macro PC1	0.89	0.84	0.64	0.71	0.48	0.63	1.00	0.17	-0.51
Macro PC2	0.47	0.42	0.38	0.41	0.15	-0.22	0.17	1.00	-0.17
Macro PC4	-0.43	-0.23	-0.22	-0.52	-0.33	-0.59	-0.51	-0.17	1.00
	Panel B. Number of factors = 7								
GDP growth	1.00	0.90	0.71	0.67	0.50	0.34	0.89	0.47	-0.43
IP growth	0.90	1.00	0.74	0.70	0.40	0.33	0.84	0.43	-0.25
Durable	0.71	0.74	1.00	0.68	0.37	0.30	0.65	0.39	-0.23
Nondurable	0.67	0.70	0.68	1.00	0.43	0.43	0.68	0.39	-0.52
Service	0.50	0.40	0.37	0.43	1.00	0.18	0.52	0.14	-0.34
Dividend growth	0.34	0.33	0.30	0.43	0.18	1.00	0.61	-0.26	-0.55
Macro PC1	0.89	0.84	0.65	0.68	0.52	0.61	1.00	0.16	-0.52
Macro PC2	0.47	0.43	0.39	0.39	0.14	-0.26	0.16	1.00	-0.17
Macro PC4	-0.43	-0.25	-0.23	-0.52	-0.34	-0.55	-0.52	-0.17	1.00

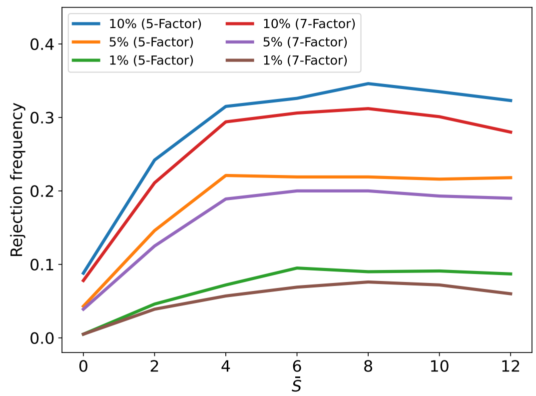
The table reports the correlation among the moving average components spanned by asset returns' latent factors, $\sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top \boldsymbol{v}_{t-s}$, with $\bar{S} = 12$ quarters. The cross-section of test assets consists of FF275. We consider six- and seven-factor models for asset returns. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

Table IA.XIV: Can external variables predict principal components of asset returns?

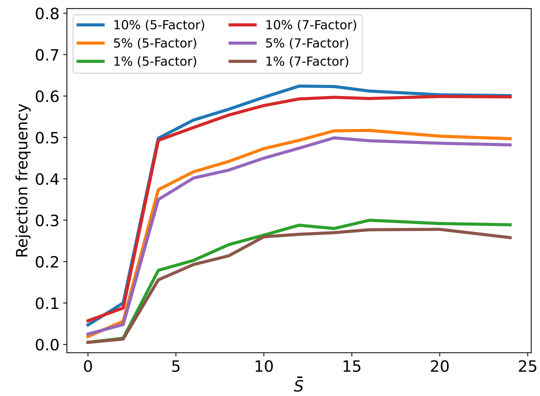
	Quarterly: Q3 1963 – Q4 2019					Monthly: July 1963 – Dec 2019				
	$PC1_t$	$PC2_t$	$PC3_t$	$PC4_t$	$PC5_t$	$PC1_t$	$PC2_t$	$PC3_t$	$PC4_t$	$PC5_t$
Panel A. Use only external predictors										
PE_{t-1}	0.058	-0.001	0.025	0.122	0.021	0.032	0.000	-0.002	-0.035	0.065
	(0.08)	(0.081)	(0.081)	(0.08)	(0.081)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
TS_{t-1}	-0.056	0.054	0.085	0.095	0.048	-0.037	0.023	-0.053	-0.043	0.05
	(0.071)	(0.072)	(0.072)	(0.071)	(0.072)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
DS_{t-1}	-0.077	-0.066	0.053	-0.092	-0.012	-0.054	-0.029	-0.032	0.038	-0.055
	(0.08)	(0.081)	(0.08)	(0.079)	(0.081)	(0.046)	(0.046)	(0.046)	(0.046)	(0.046)
VS_{t-1}	-0.109	-0.020	-0.009	-0.030	-0.050	-0.088**	-0.017	0.028	0.060	0.005
	(0.073)	(0.074)	(0.074)	(0.073)	(0.074)	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
R^2_{adj}	0.92%	-1.31%	-0.72%	1.58%	-1.49%	0.79%	-0.49%	-0.08%	-0.07%	0.57%
Panel B. Use both external predictors and lagged PCs										
$PC1_{t-1}$	-0.04	0.097	-0.069	-0.305***	0.036	0.128***	0.227***	0.075*	-0.121***	-0.162***
	(0.067)	(0.067)	(0.066)	(0.064)	(0.068)	(0.038)	(0.038)	(0.039)	(0.038)	(0.037)
$PC2_{t-1}$	-0.154**	0.024	0.042	0.059	-0.161**	-0.05	-0.02	-0.041	0.023	-0.189***
	(0.066)	(0.066)	(0.065)	(0.063)	(0.067)	(0.038)	(0.038)	(0.038)	(0.038)	(0.037)
$PC3_{t-1}$	0.061	-0.114*	0.253***	-0.028	0.053	-0.021	-0.051	0.118***	0.05	0.008
	(0.068)	(0.068)	(0.067)	(0.065)	(0.069)	(0.038)	(0.038)	(0.039)	(0.039)	(0.037)
$PC4_{t-1}$	0.005	0	-0.174***	-0.047	0.044	-0.008	0.006	-0.043	-0.078**	-0.12***
	(0.067)	(0.067)	(0.066)	(0.064)	(0.068)	(0.038)	(0.038)	(0.038)	(0.038)	(0.037)
$PC5_{t-1}$	0.13*	-0.226***	-0.052	0.14**	0.075	0.09**	0.025	0.02	-0.075**	0.004
	(0.067)	(0.067)	(0.066)	(0.065)	(0.068)	(0.038)	(0.038)	(0.038)	(0.038)	(0.037)
PE_{t-1}	0.063	0.013	0.025	0.103	0.029	0.029	0.002	0.004	-0.032	0.067
	(0.08)	(0.08)	(0.079)	(0.077)	(0.081)	(0.046)	(0.045)	(0.046)	(0.046)	(0.045)
TS_{t-1}	-0.065	0.066	0.012	0.068	0.049	-0.036	0.03	-0.039	-0.041	0.044
	(0.073)	(0.073)	(0.071)	(0.07)	(0.074)	(0.041)	(0.04)	(0.041)	(0.041)	(0.04)
DS_{t-1}	-0.094	-0.021	0.055	-0.125	-0.016	-0.039	-0.015	-0.021	0.027	-0.061
	(0.08)	(0.08)	(0.078)	(0.077)	(0.081)	(0.046)	(0.045)	(0.046)	(0.046)	(0.044)
VS_{t-1}	-0.113	-0.003	0.064	-0.084	-0.042	-0.07*	0.016	0.017	0.036	-0.025
	(0.075)	(0.075)	(0.074)	(0.072)	(0.076)	(0.042)	(0.042)	(0.042)	(0.042)	(0.041)
R^2_{adj}	3.34%	3.78%	6.88%	11.12%	0.05%	2.8%	4.21%	1.52%	2.11%	7.51%

The table reports the empirical results of regressing principal components (PCs) of asset returns on their one-period lags and external predictors. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider the predictability of the five largest PCs at both quarterly (left side) and monthly (right side) frequencies. External predictors include the price-earning ratio of the SP500 index (PE_{t-1}), term spread (TS_{t-1}), default spread (DS_{t-1}), and value spread (VS_{t-1}). The numbers without parentheses are coefficient estimates, with their standard errors in the parentheses. If the coefficient estimate is significant in the 90% (95%, 99%) significance level, it will be highlighted by * (**, ***). The final row shows the adjusted R-squared (R^2_{adj}) in each regression. In Panel A, we regress PCs on only the four external predictors, while we further add the one-period lags of PCs in Panel B. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

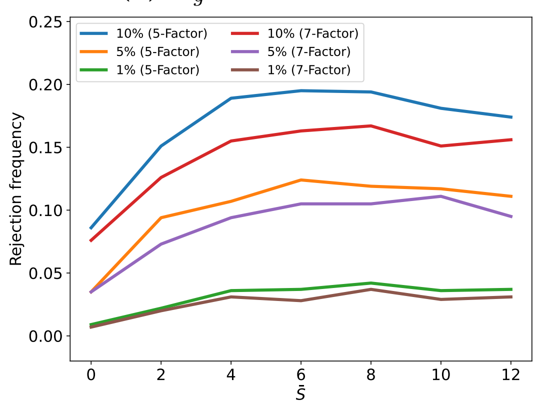
IA.7 Additional Figures



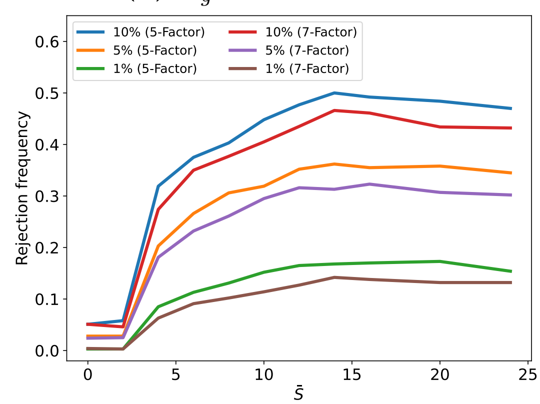
(a) $R_g^2 = 20\%$ & $T = 200$



(b) $R_g^2 = 20\%$ & $T = 600$



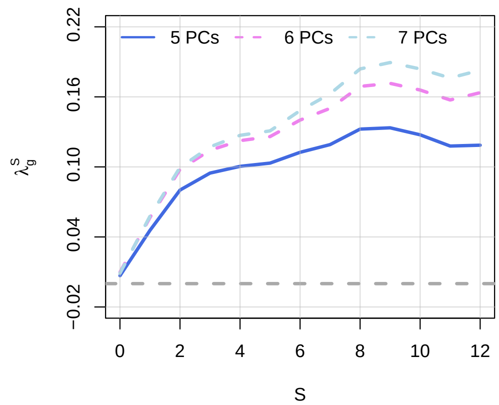
(c) $R_g^2 = 10\%$ & $T = 200$



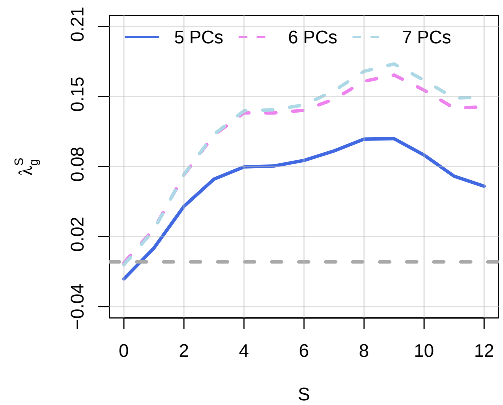
(d) $R_g^2 = 10\%$ & $T = 600$

Figure IA.1: Power of identifying strong factors

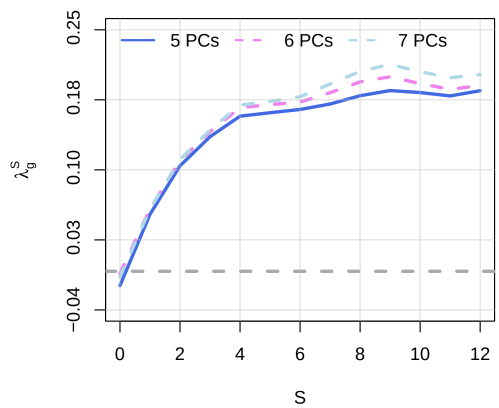
The figure plots the frequency of rejecting the null hypothesis $H_0 : \lambda_g^{\bar{S}} = 0$ based on the 90%, 95%, and 99% credible intervals based on our Bayesian estimates in Proposition 1. $\lambda_g^{\bar{S}}$ is defined in equation (6). We consider strong factors, with $R_g^2 \in \{10\%, 20\%, 30\%\}$, and two sample sizes, $T \in \{200, 600\}$. In each simulated scenario, we estimate several model configurations with different numbers of factors and different $\bar{S}$. The number of Monte Carlo simulations is 1,000.



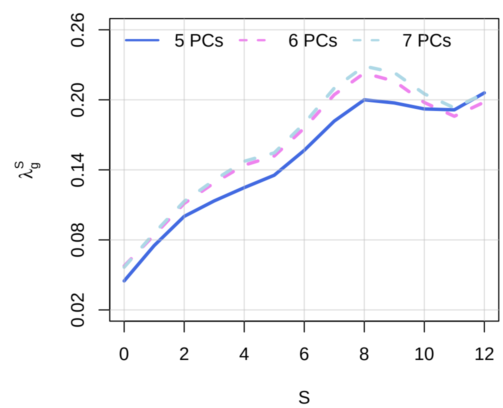
(a) GDP growth



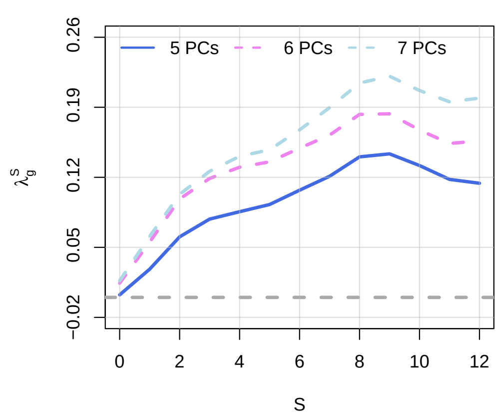
(b) IP growth



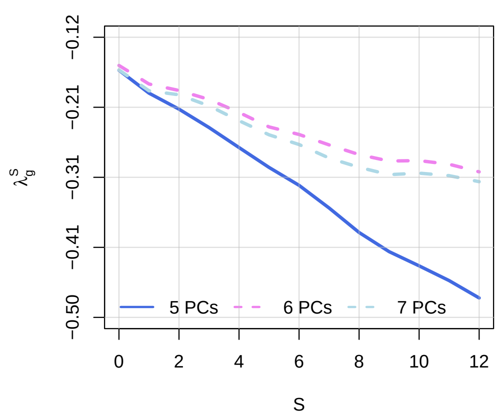
(c) Durable consumption growth



(d) Nondurable consumption growth



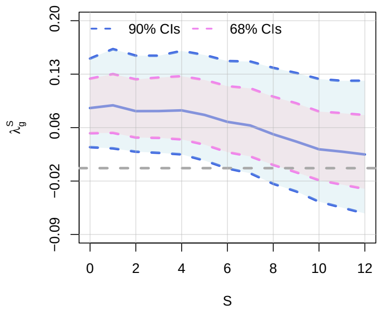
(e) Macro PC1 (FRED-QD)



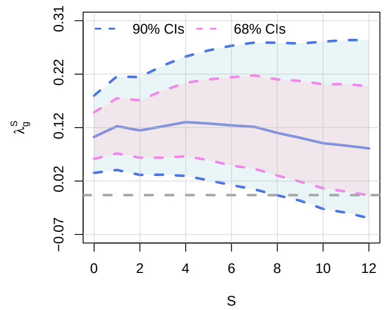
(f) Macro PC4 (FRED-QD)

Figure IA.2: Per-period mean returns of *horizon-specific* mimicking portfolios

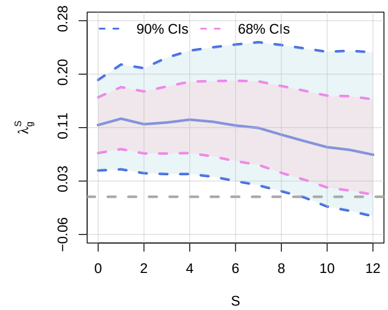
This figure plots the per-period mean returns of $g_{t-1\to t+S}$'s horizon-specific mimicking portfolio, where S ranges from 0 to 12 quarters. In particular, we project nontraded risk factor onto PCs of asset returns across different horizons: $\boldsymbol{\omega}_S^{MP} = \text{cov}(\boldsymbol{v}_{t-1\to t+S})^{-1}\text{cov}(\boldsymbol{v}_{t-1\to t+S}, g_{t-1\to t+S})$, where $\boldsymbol{v}_{t-1\to t+S}$ are the cumulative returns on the PCs between $t-1$ and $t+S$. Next, we estimate time-series averages of $(\boldsymbol{\omega}_S^{MP})^\top \boldsymbol{v}_t$, $0 \le S \le 12$. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider the first five, six, and seven PCs in constructing the horizon-specific mimicking portfolios.



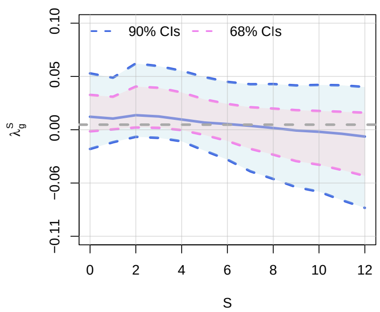
(a) AEM: 5 factors



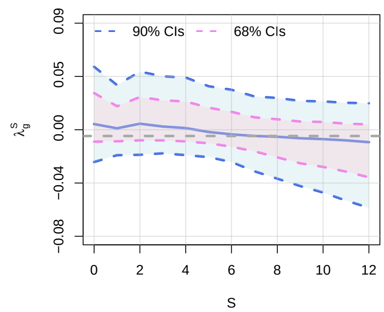
(b) AEM: 6 factors



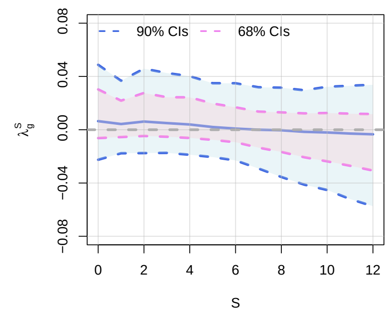
(c) AEM: 7 factors



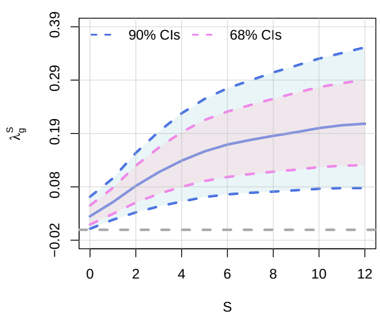
(d) Capital Share: 5 factors



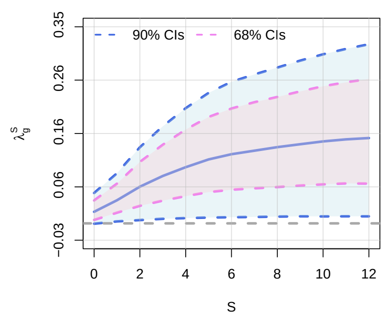
(e) Capital Share: 6 factors



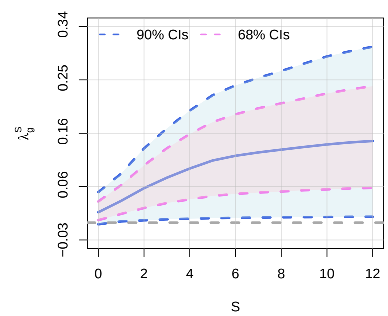
(f) Capital Share: 7 factors



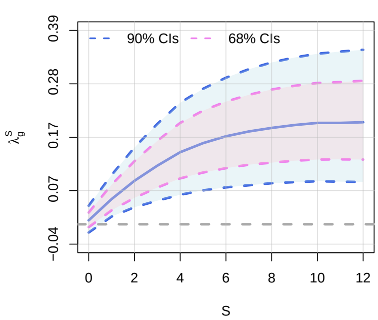
(g) GDP: 5 factors



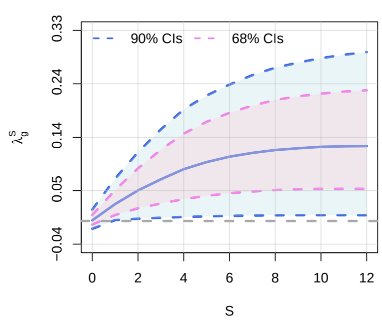
(h) GDP: 6 factors



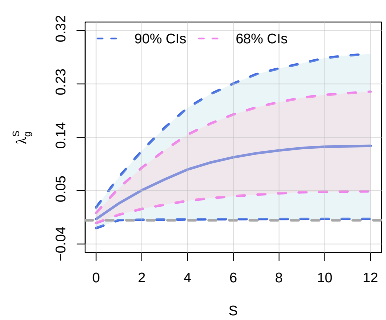
(i) GDP: 7 factors



(j) IP growth: 5 factors

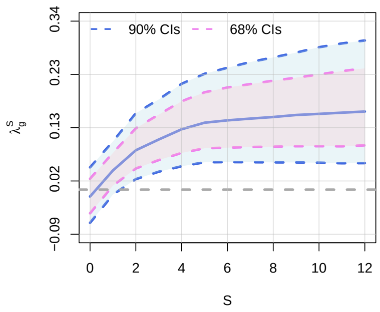


(k) IP growth: 6 factors

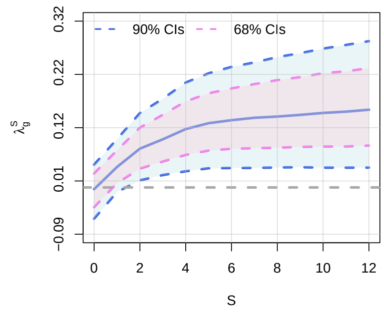


(l) IP growth: 7 factors

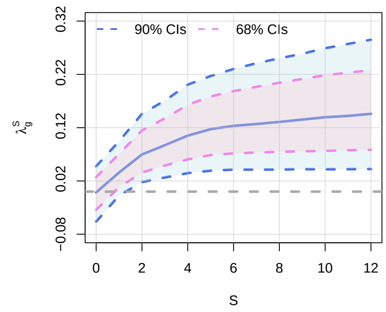
Figure IA.3: Term structure of factor's risk premia: Quarterly variables



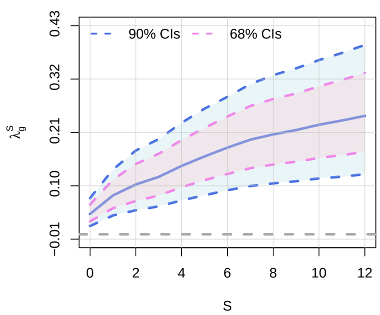
(m) Durable: 5 factors



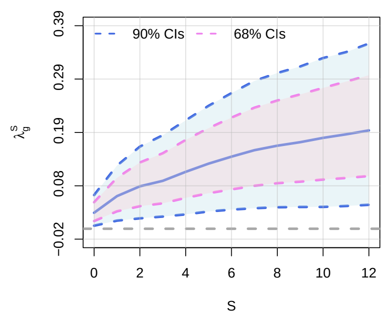
(n) Durable: 6 factors



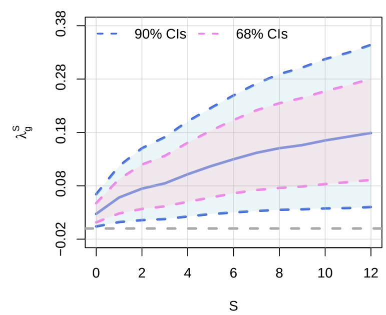
(o) Durable: 7 factors



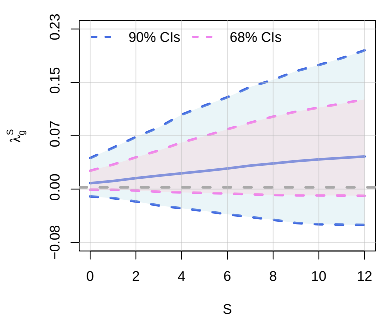
(p) Nondurable: 5 factors



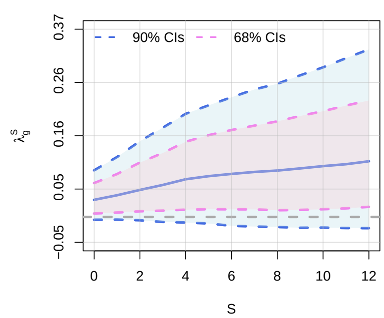
(q) Nondurable: 6 factors



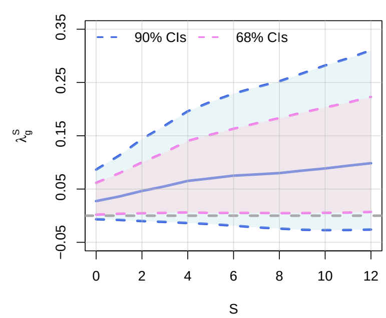
(r) Nondurable: 7 factors



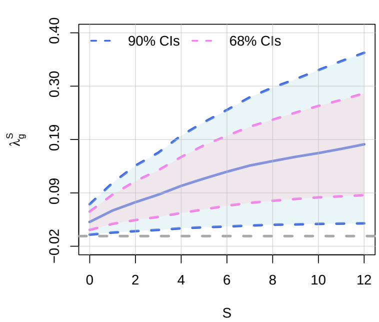
(s) Service: 5 factors



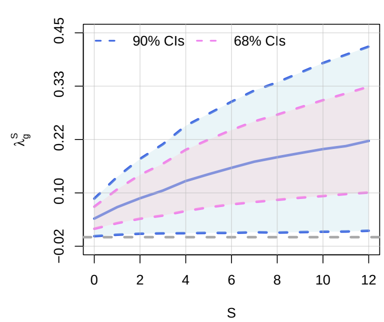
(t) Service: 6 factors



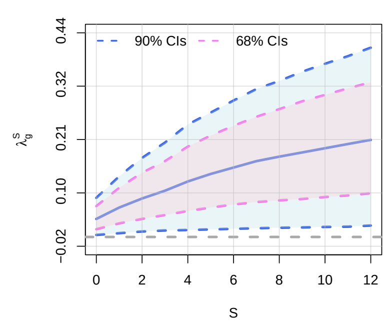
(u) Service: 7 factors



(v) Nondur+service: 5 factors

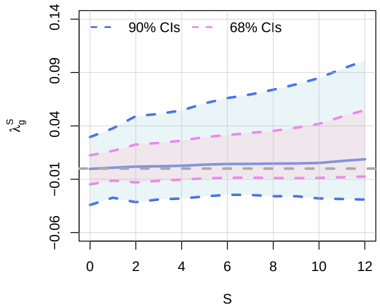


(w) Nondur+Service: 6 factors

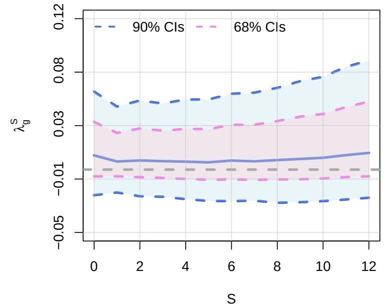


(x) Nondur+service: 7 factors

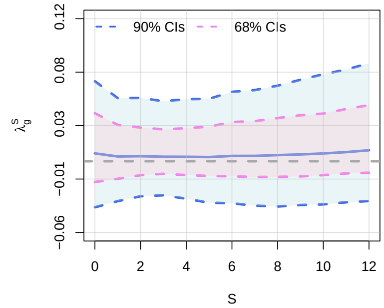
Figure IA.3: Term structure of factor's risk premia: Quarterly variables



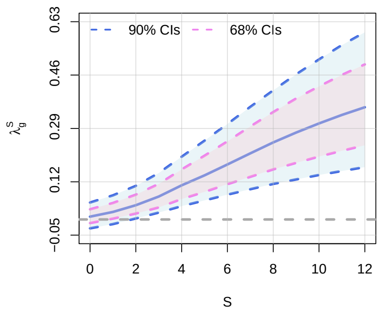
(y) Labor: 5 factors



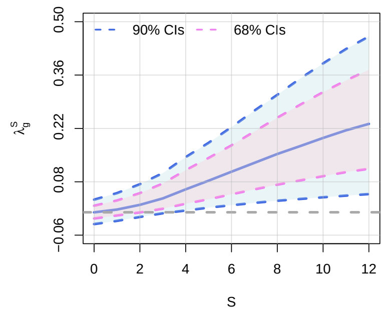
(z) Labor: 6 factors



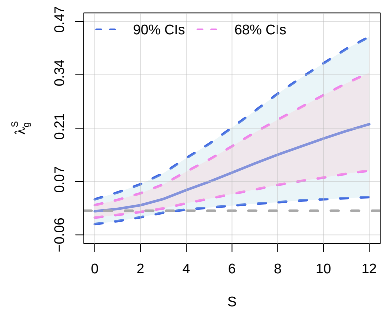
(aa) Labor: 7 factors



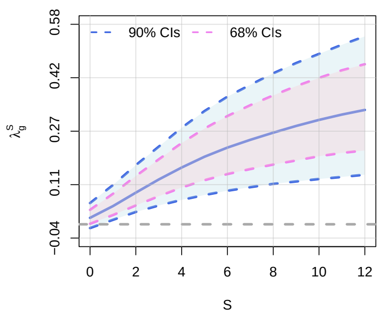
(ab) Dividend: 5 factors



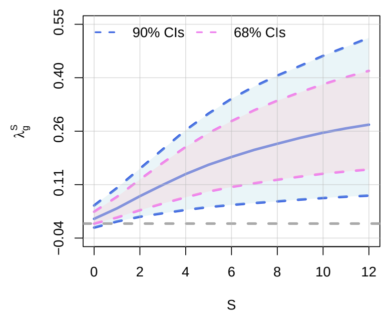
(ac) Dividend: 6 factors



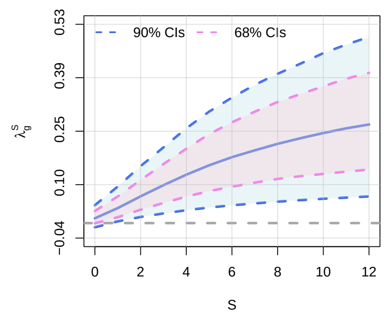
(ad) Dividend: 7 factors



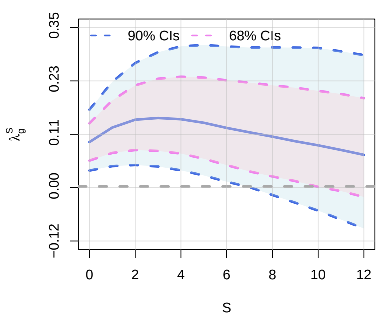
(ae) Macro PC1: 5 factors



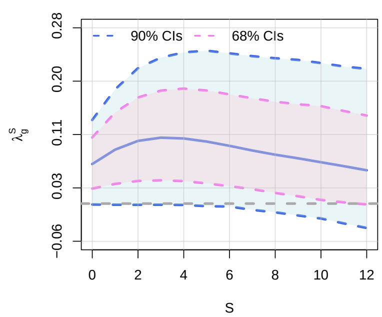
(af) Macro PC1: 6 factors



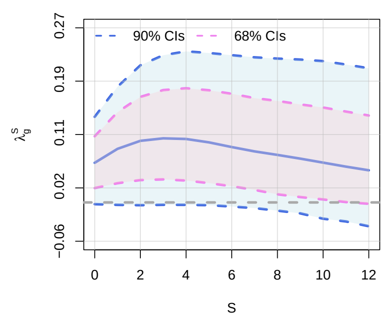
(ag) Macro PC1: 7 factors



(ah) Macro PC2: 5 factors

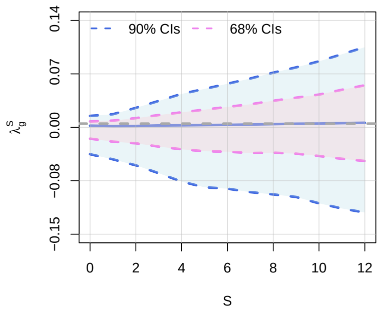


(ai) Macro PC2: 6 factors

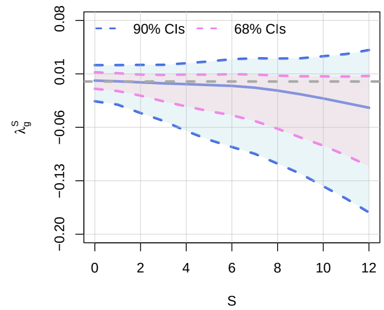


(aj) Macro PC2: 7 factors

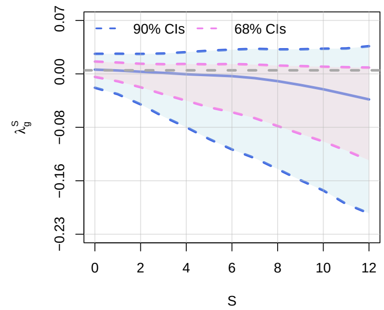
Figure IA.3: Term structure of factor's risk premia: Quarterly variables



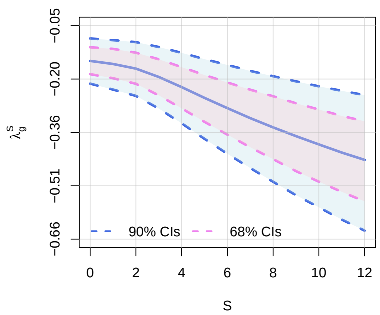
(ak) Macro PC3: 5 factors



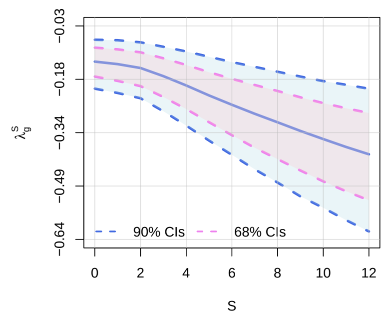
(al) Macro PC3: 6 factors



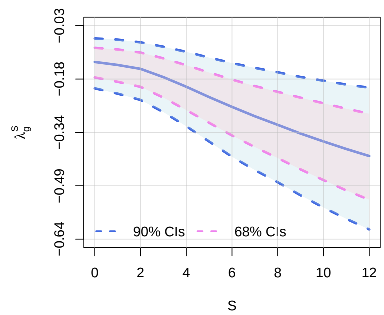
(am) Macro PC3: 7 factors



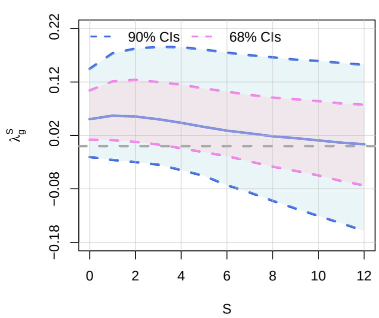
(an) Macro PC4: 5 factors



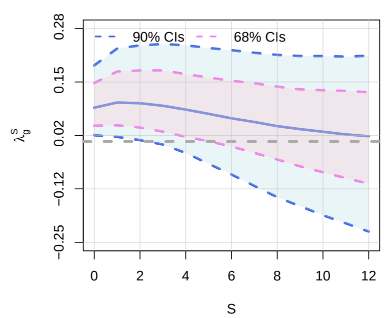
(ao) Macro PC4: 6 factors



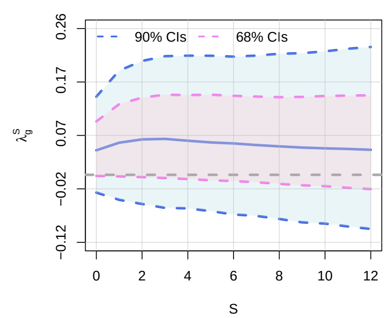
(ap) Macro PC4: 7 factors



(aq) Macro PC5: 5 factors



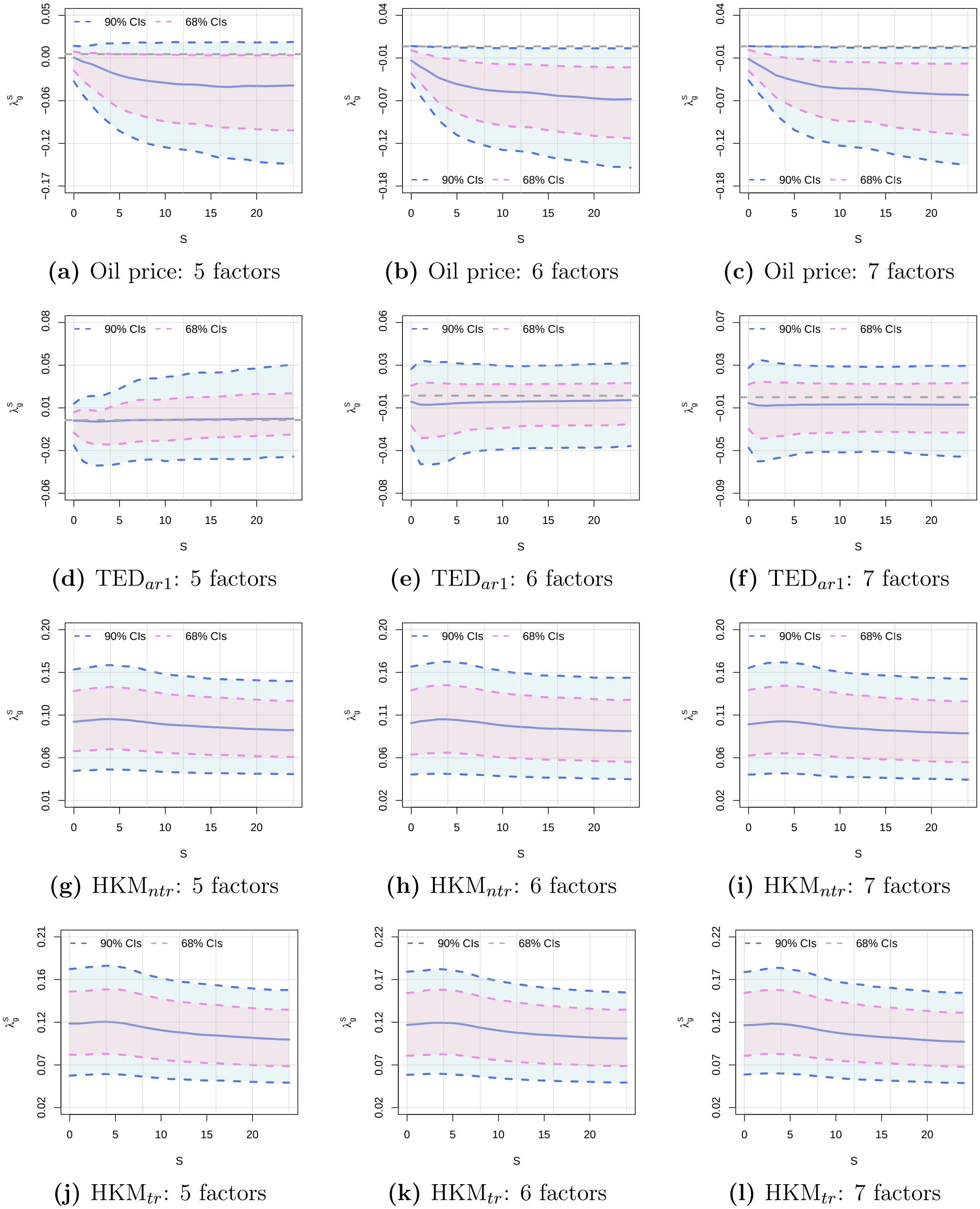
(ar) Macro PC5: 6 factors



(as) Macro PC5: 7 factors

Figure IA.3: Term structure of factor's risk premia: Quarterly variables

The figure plots the term structure of risk premia estimates using Proposition 1, where the risk premia over S horizons (λ_g^S) are defined in equation (6). The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-, six- and seven-factor models for asset returns. We study quarterly factors, whose risk premia are estimated using a lag of 12 quarters in g_t's equations. In addition to the point estimates, we show the 68% and 90% Bayesian credible intervals, highlighted in pink and blue, respectively. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.



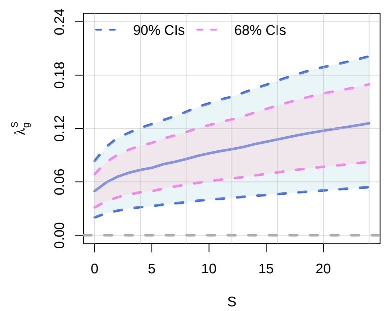
(a) Oil price: 5 factors (b) Oil price: 6 factors (c) Oil price: 7 factors

(d) TED$_{ar1}$: 5 factors (e) TED$_{ar1}$: 6 factors (f) TED$_{ar1}$: 7 factors

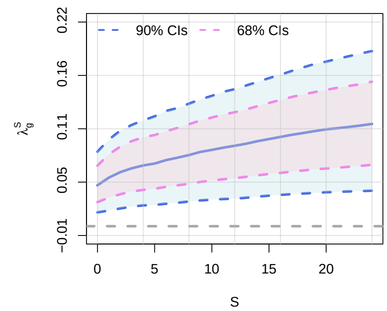
(g) HKM$_{ntr}$: 5 factors (h) HKM$_{ntr}$: 6 factors (i) HKM$_{ntr}$: 7 factors

(j) HKM$_{tr}$: 5 factors (k) HKM$_{tr}$: 6 factors (l) HKM$_{tr}$: 7 factors

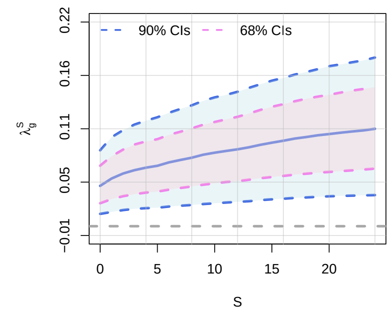
Figure IA.4: Term structure of factor's risk premia: Monthly variables



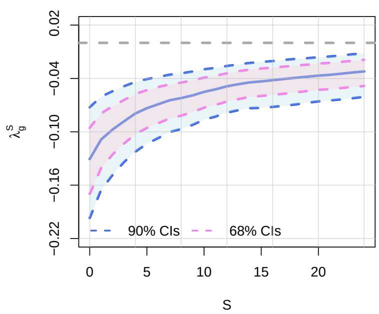
(m) PS liquidity: 5 factors



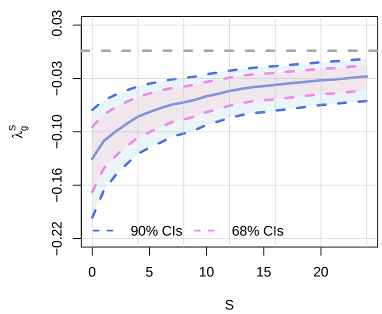
(n) PS liquidity: 6 factors



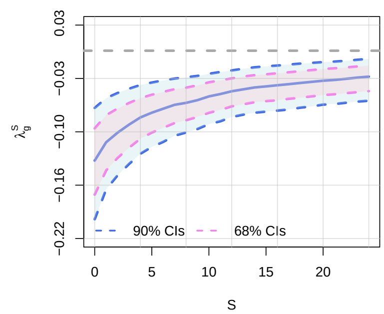
(o) PS liquidity: 7 factors



(p) $\Delta \log(\text{VIX})$: 5 factors



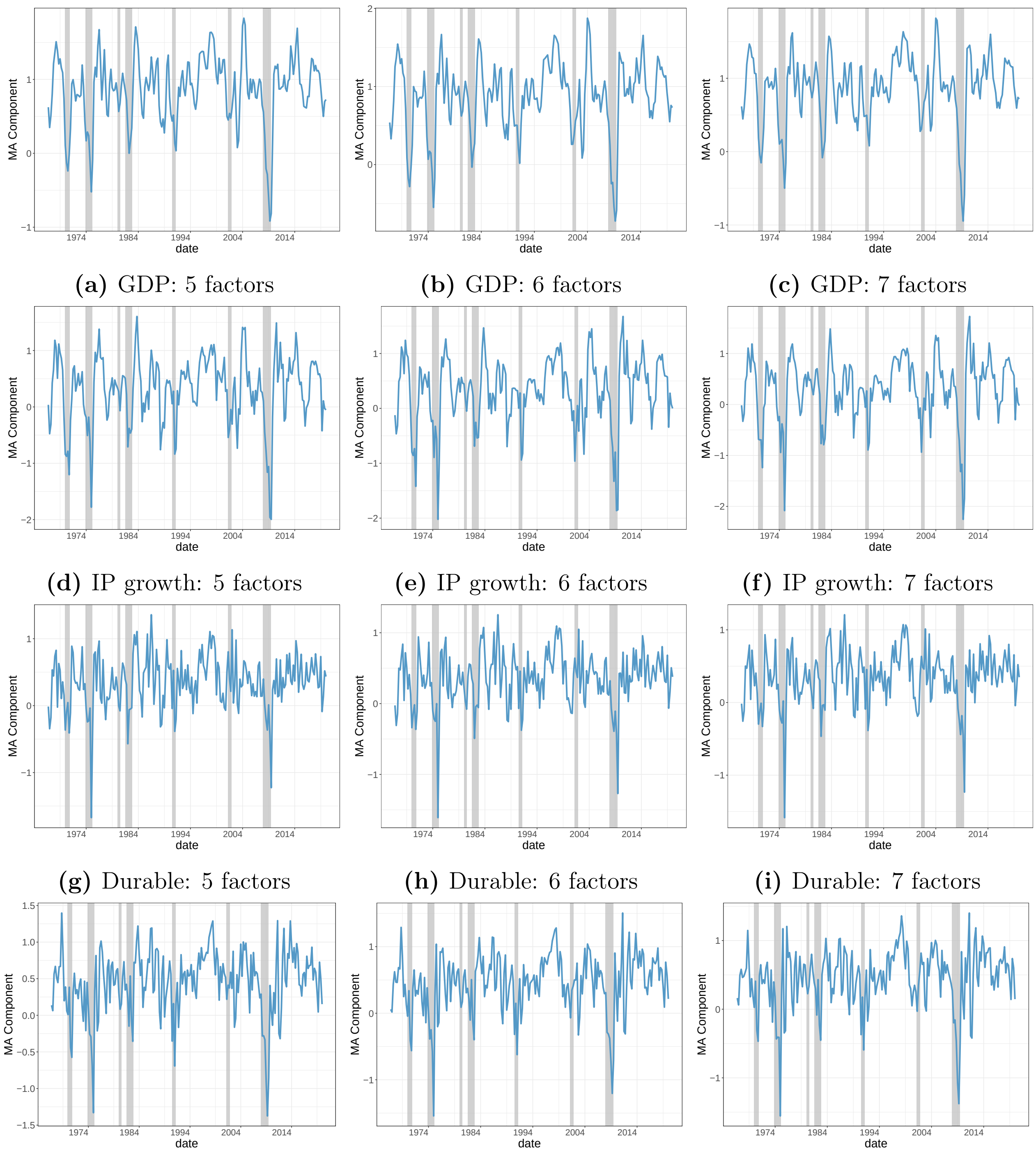
(q) $\Delta \log(\text{VIX})$: 6 factors



(r) $\Delta \log(\text{VIX})$: 7 factors

Figure IA.4: Term structure of factor's risk premia: Monthly variables

The figure plots the term structure of risk premia estimates using Proposition 1, where the risk premia over S horizons (λ_g^S) are defined in equation (6). The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-, six- and seven-factor models for asset returns. We study monthly factors, whose risk premia are estimated using a lag of 24 months in g_t's equations. For Fama-French five factors, we also include their in-sample monthly Sharpe ratios (see black dotted lines). In addition to the point estimates, we show the 68% and 90% Bayesian credible intervals, highlighted in pink and blue, respectively. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.



(a) GDP: 5 factors

(b) GDP: 6 factors

(c) GDP: 7 factors

(d) IP growth: 5 factors

(e) IP growth: 6 factors

(f) IP growth: 7 factors

(g) Durable: 5 factors

(h) Durable: 6 factors

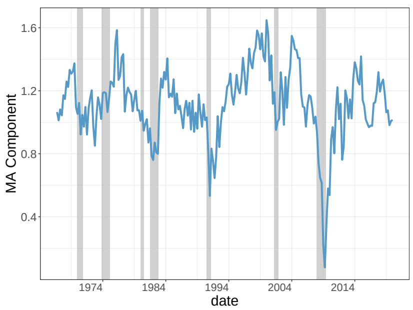
(i) Durable: 7 factors

(j) Nondurable: 5 factors

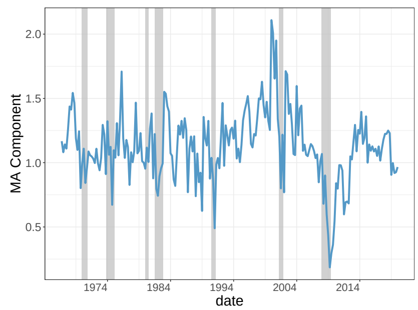
(k) Nondurable: 6 factors

(l) Nondurable: 7 factors

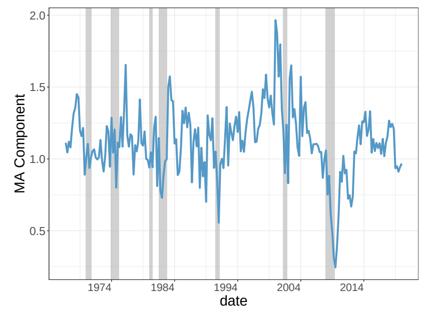
Figure IA.5: Moving average components of some macro factors



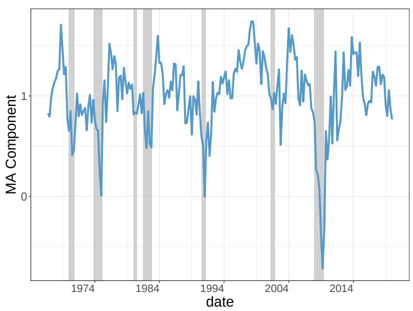
(m) Service: 5 factors



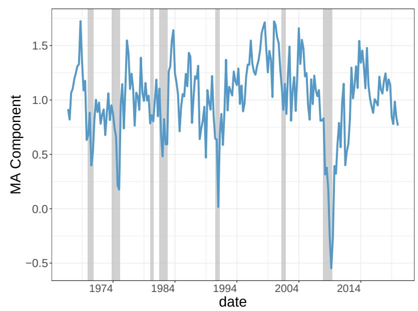
(n) Service: 6 factors



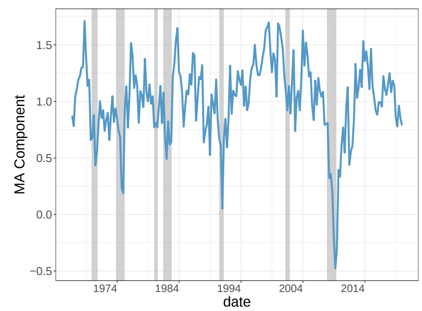
(o) Service: 7 factors



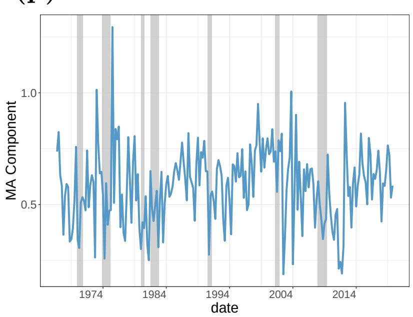
(p) Nondur+service: 5 factors



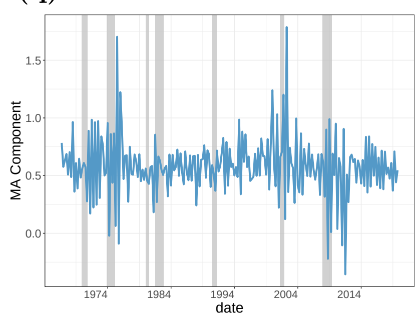
(q) Nondur+Service: 6 factors



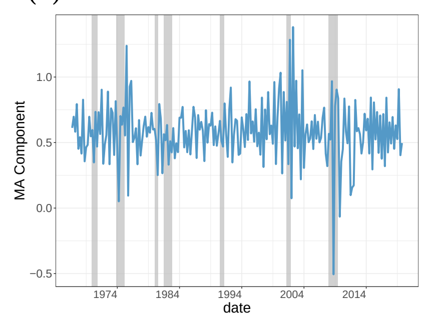
(r) Nondur+service: 7 factors



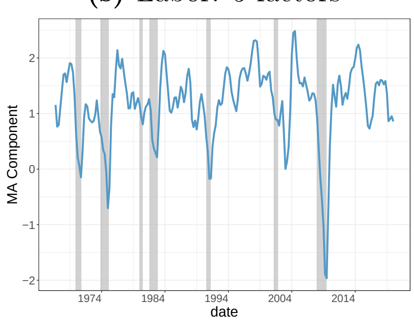
(s) Labor: 5 factors



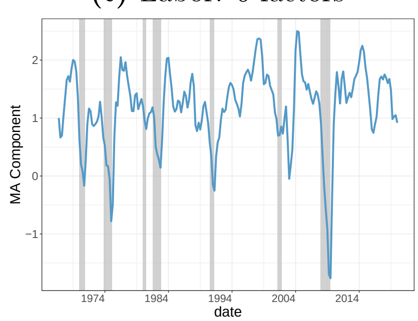
(t) Labor: 6 factors



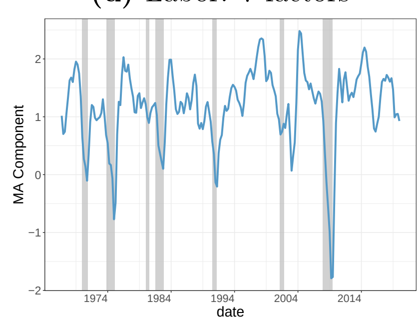
(u) Labor: 7 factors



(v) Macro PC1: 5 factors

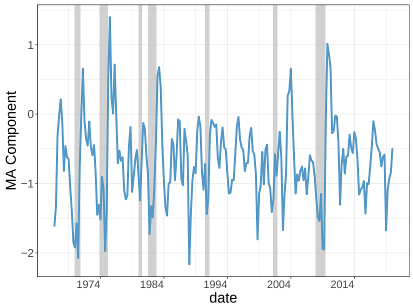


(w) Macro PC1: 6 factors

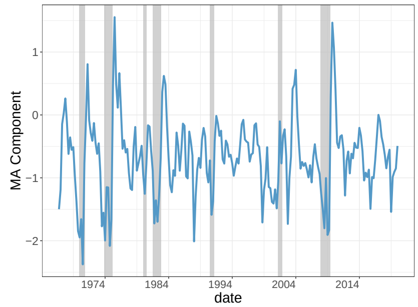


(x) Macro PC1: 7 factors

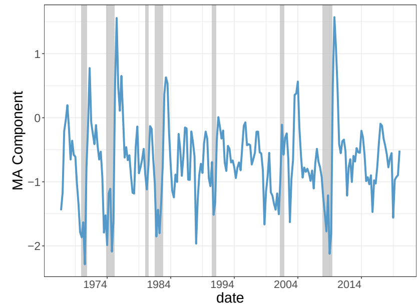
Figure IA.5: Moving average components of some macro factors



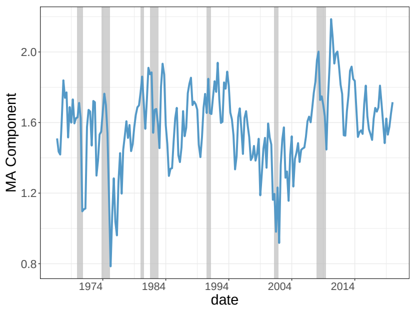
(y) Macro PC2: 5 factors



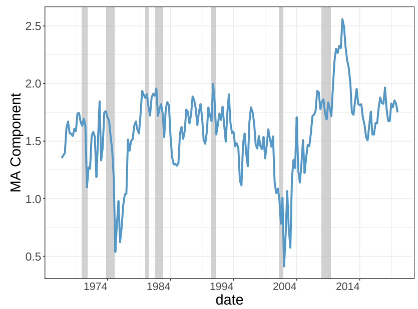
(z) Macro PC2: 6 factors



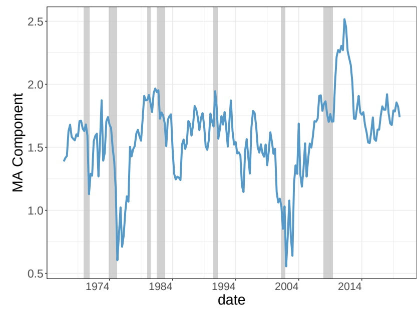
(aa) Macro PC2: 7 factors



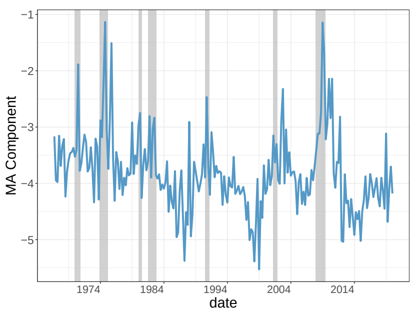
(ab) Macro PC3: 5 factors



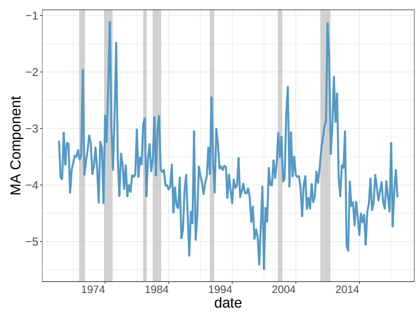
(ac) Macro PC3: 6 factors



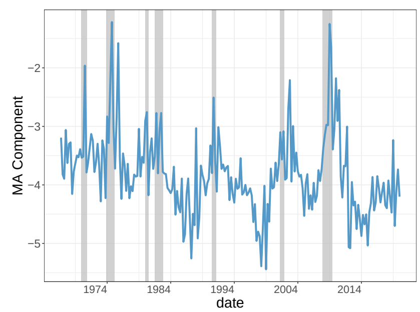
(ad) Macro PC3: 7 factors



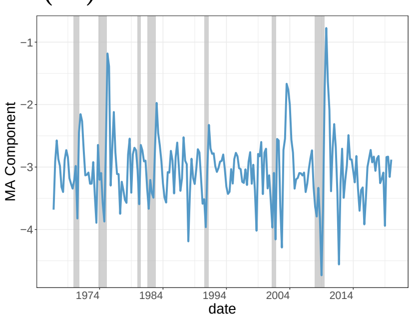
(ae) Macro PC4: 5 factors



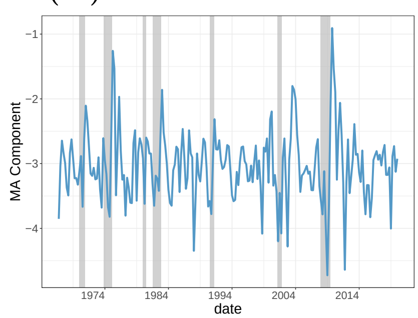
(af) Macro PC4: 6 factors



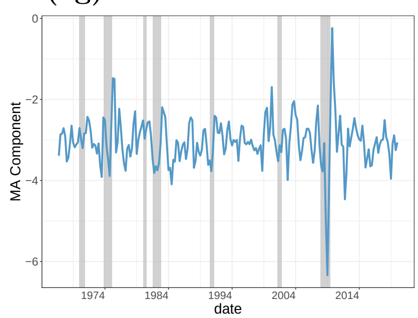
(ag) Macro PC4: 7 factors



(ah) Macro PC5: 5 factors



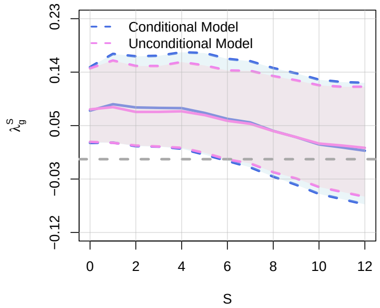
(ai) Macro PC5: 6 factors



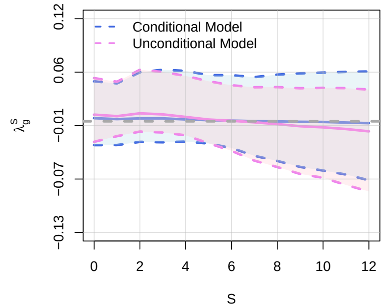
(aj) Macro PC5: 7 factors

Figure IA.5: Moving average components of some macro factors

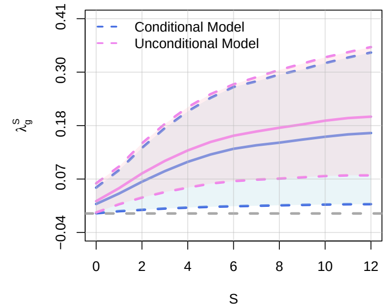
The figure plots the time series of (posterior means of) moving average components spanned by asset returns' latent factors: $\sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top \boldsymbol{v}_{t-s}$, with $\bar{S} = 12$ quarters. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider five-, six- and seven-factor models for asset returns. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.



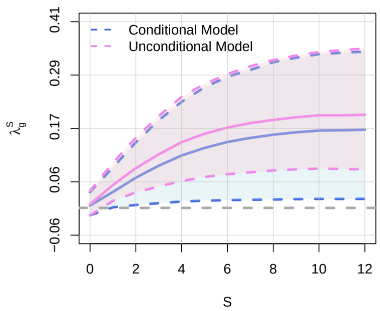
(a) AEM intermediary



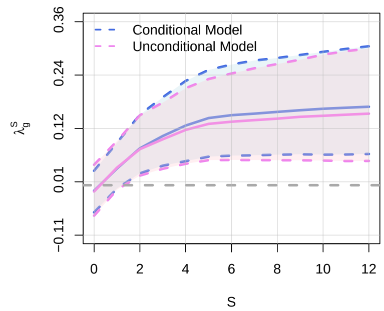
(b) Capital share



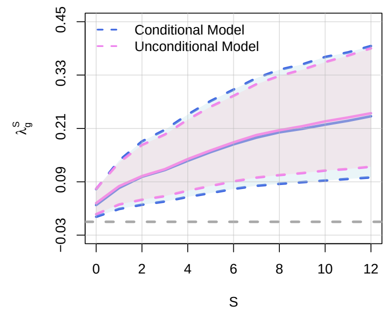
(c) GDP growth



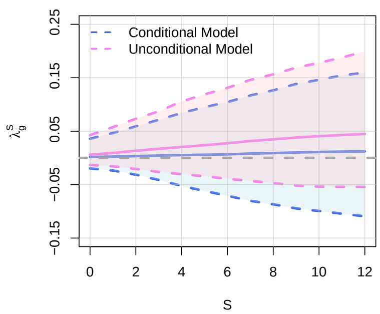
(d) IP growth



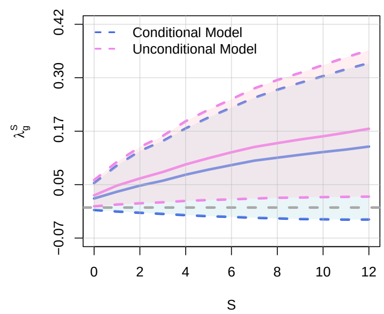
(e) Durable cons



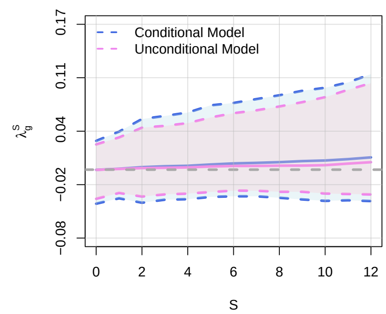
(f) Nondurable cons



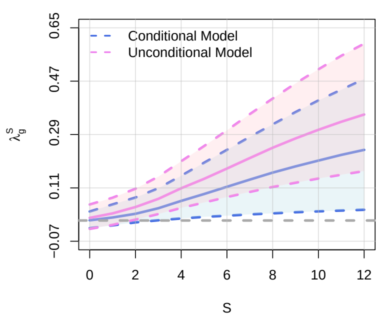
(g) Service cons



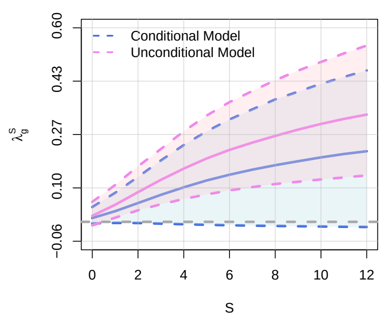
(h) Nondurable+service



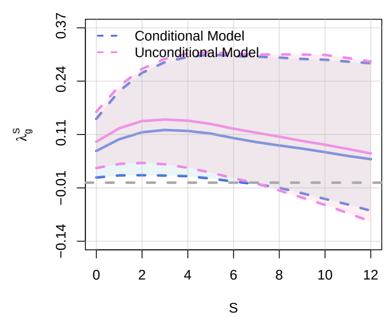
(i) Labor income



(j) Dividend growth

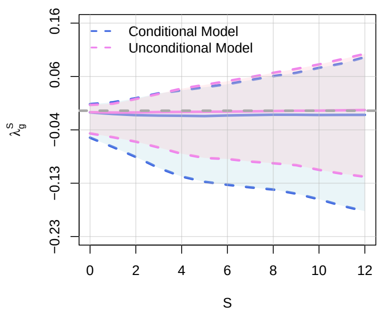


(k) Macro PC1

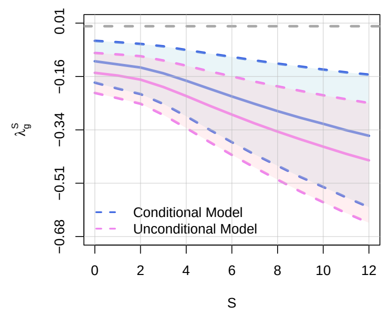


(l) Macro PC2

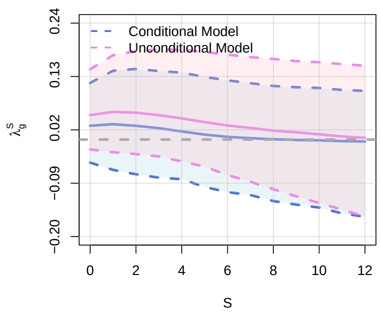
Figure IA.6: Term structure of unconditional risk premia in time-varying models



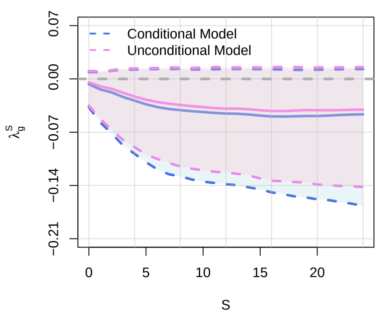
(m) Macro PC3



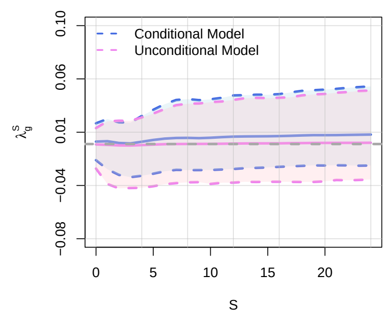
(n) Macro PC4



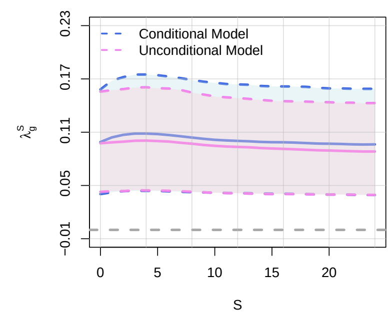
(o) Macro PC5



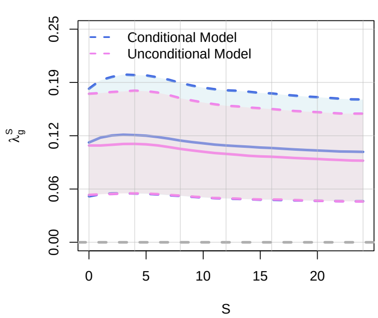
(p) Oil price change



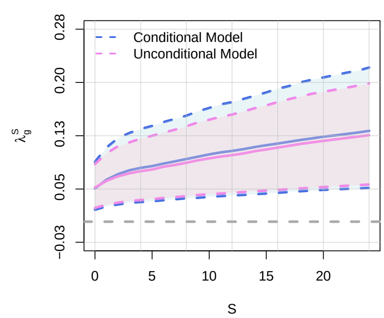
(q) TED spread



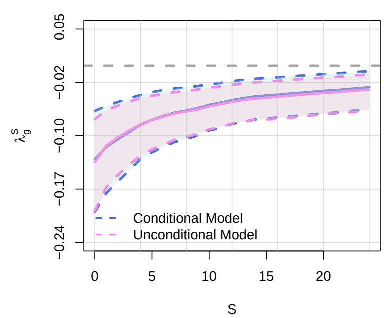
(r) HKM_{ntr}



(s) HKM_{tr}



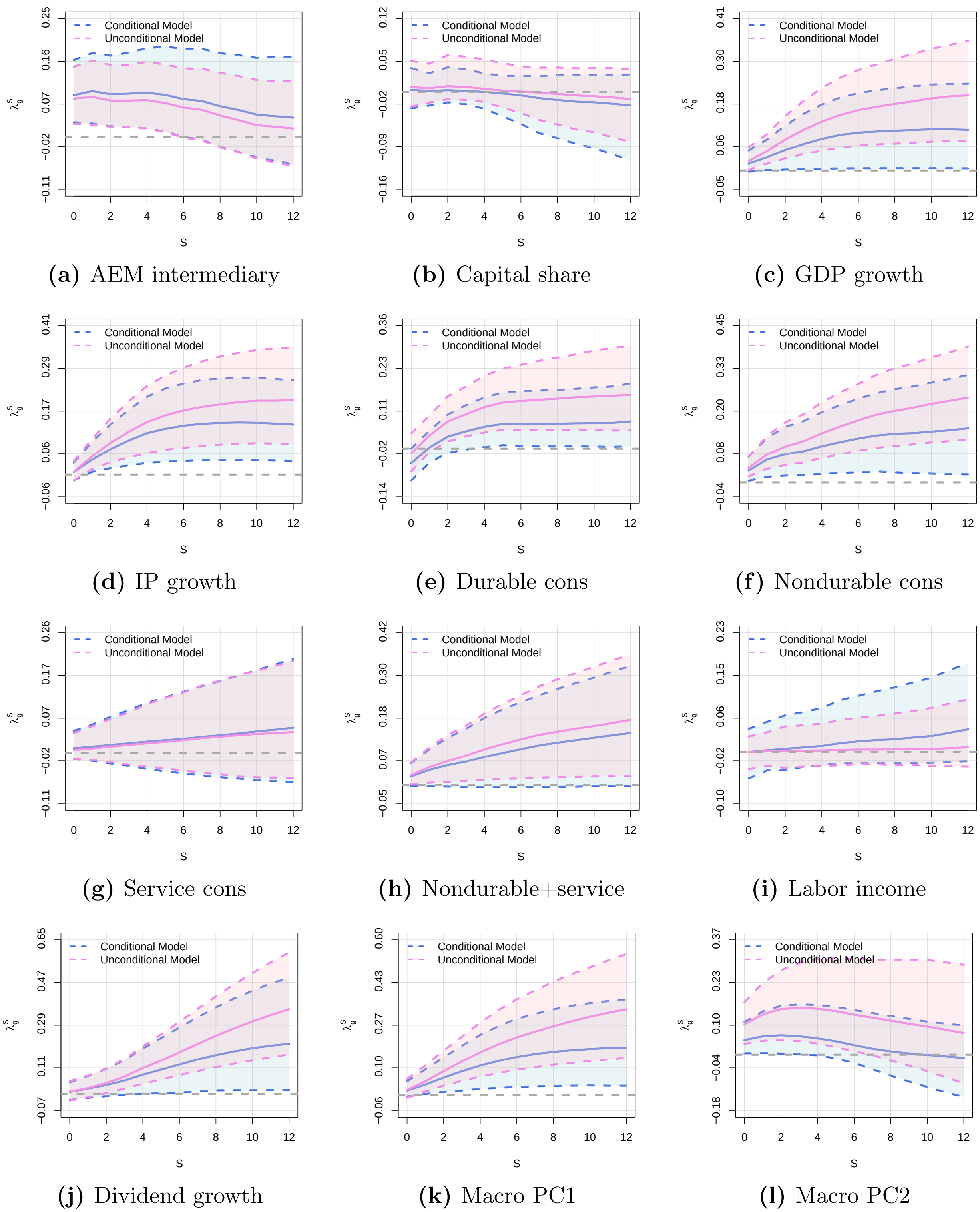
(t) PS liquidity



(u) $\Delta\log(\text{VIX})$

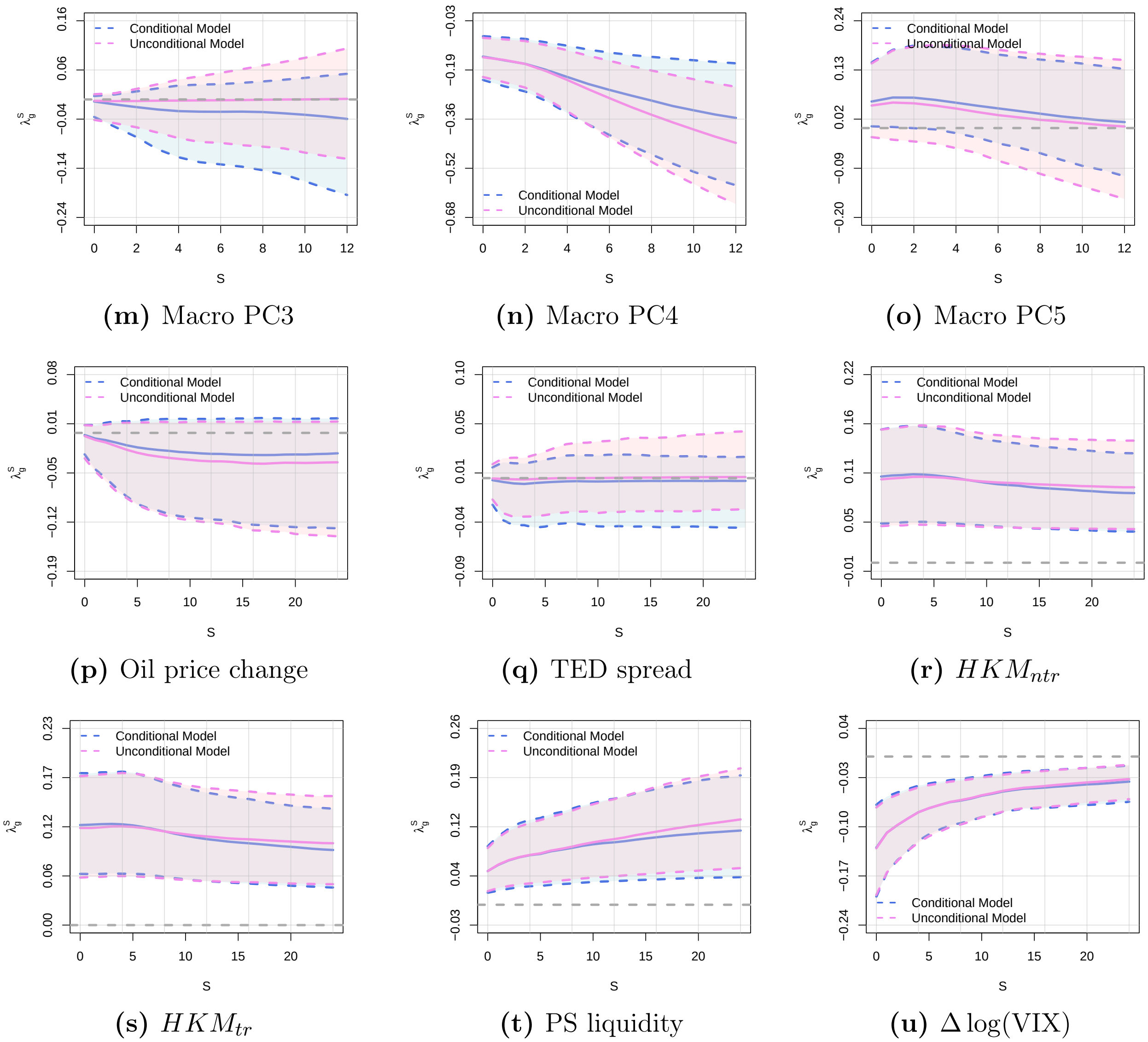
Figure IA.6: Term structure of unconditional risk premia in time-varying models

The figure plots the term structure of unconditional risk premia estimates using Propositions 1 and A2. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider the five-factor models for asset returns. For quarterly (monthly) factors, risk premia are estimated using a lag of 12 quarters (24 months) in g_t's equations. The blue dotted lines and the light blue shaded areas present the risk premia estimates and their 90% posterior credible intervals under the conditional models using the method described in Section 2.2. In the time-varying models, we model the dynamics of latent systematic factors as a VAR(1) process. For comparison, we include the purple dotted lines and the related shaded areas, showing the risk premia estimated using the unconditional models described in Section 2.1. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.



(a) AEM intermediary (b) Capital share (c) GDP growth
(d) IP growth (e) Durable cons (f) Nondurable cons
(g) Service cons (h) Nondurable+service (i) Labor income
(j) Dividend growth (k) Macro PC1 (l) Macro PC2

Figure IA.7: Term structure of unconditional risk premia in time-varying models with only external predictors: PE ratio of S&P 500, Term spread, default spread, and value spread



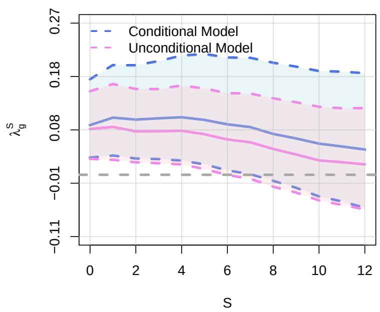
(m) Macro PC3 **(n)** Macro PC4 **(o)** Macro PC5

(p) Oil price change **(q)** TED spread **(r)** HKM_{ntr}

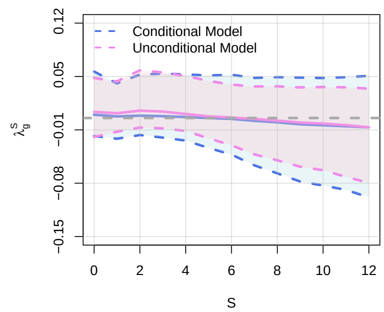
(s) HKM_{tr} **(t)** PS liquidity **(u)** $\Delta\log(\text{VIX})$

Figure IA.7: Term structure of unconditional risk premia in time-varying models with only external predictors: PE ratio of S&P 500, Term spread, default spread, and value spread

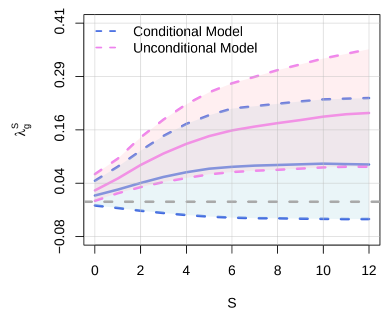
The figure plots the term structure of unconditional risk premia estimates using Propositions 1 and A2. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider the five-factor models for asset returns. For quarterly (monthly) factors, risk premia are estimated using a lag of 12 quarters (24 months) in g_t's equations. The blue dotted lines and the light blue shaded areas present the risk premia estimates and their 90% posterior credible intervals under the conditional models using the method described in Section 2.2. In the time-varying models, we model the dynamics of latent systematic factors as a VAR(1) process, where both latent factors and four external predictors (PE ratio of S&P 500, Term spread, default spread, and value spread) are driven by only the lagged external predictors. For comparison, we include the purple dotted lines and the related shaded areas, showing the risk premia estimated using the unconditional models described in Section 2.1. Definition and data sources of factors and test assets can be found in Appendix IA.3.



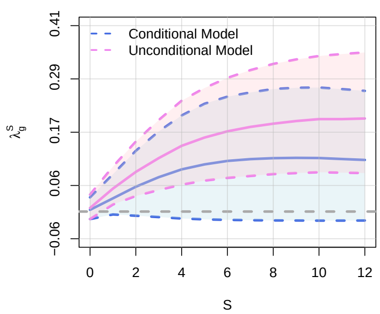
(a) AEM intermediary



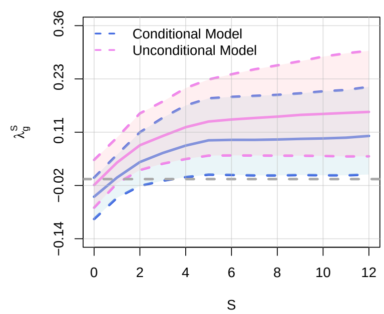
(b) Capital share



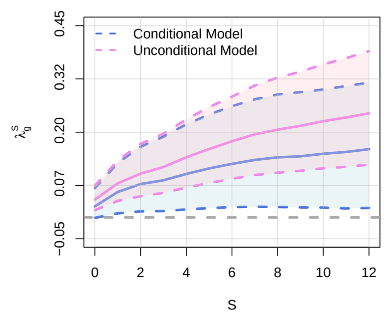
(c) GDP growth



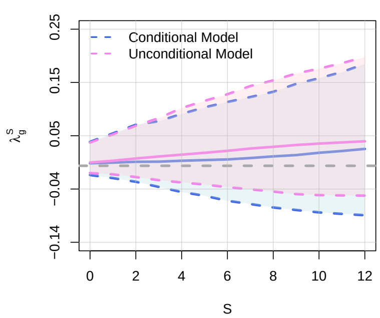
(d) IP growth



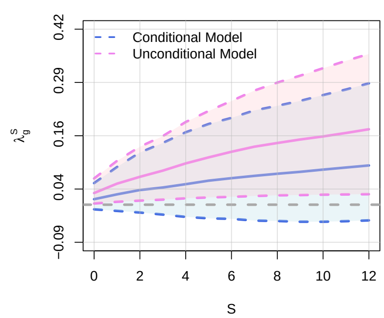
(e) Durable cons



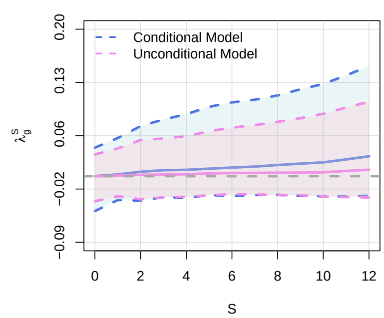
(f) Nondurable cons



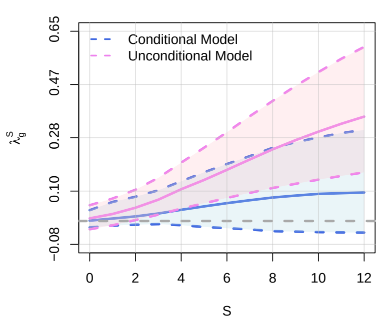
(g) Service cons



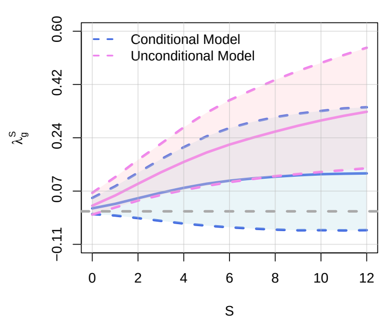
(h) Nondurable+service



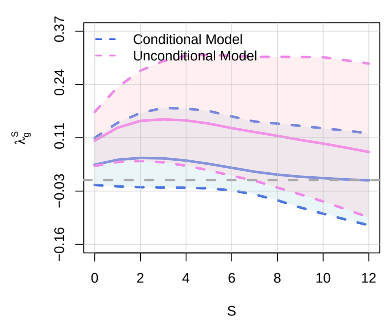
(i) Labor income



(j) Dividend growth

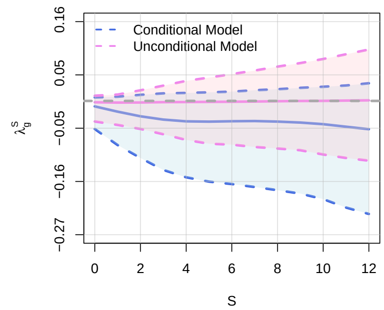


(k) Macro PC1

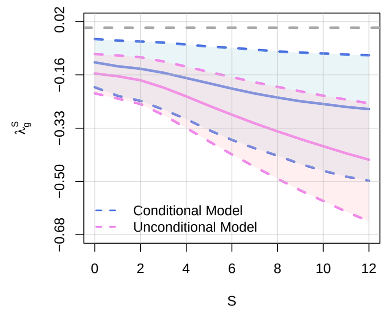


(l) Macro PC2

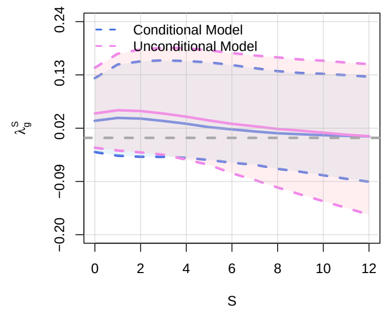
Figure IA.8: Term structure of unconditional risk premia in time-varying models with both lagged latent factors and external predictors: PE ratio of S&P 500, Term spread, default spread, and value spread



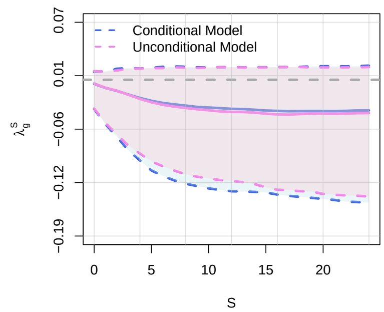
(m) Macro PC3



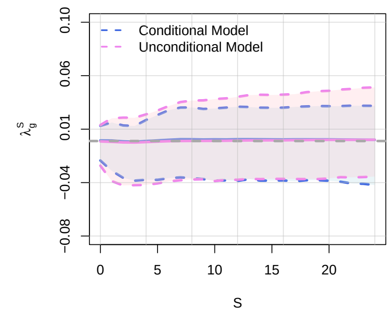
(n) Macro PC4



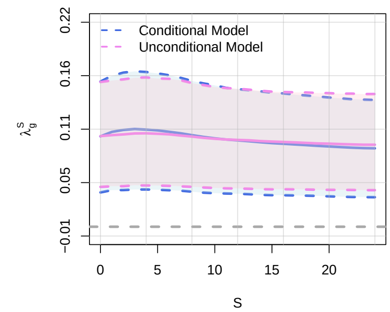
(o) Macro PC5



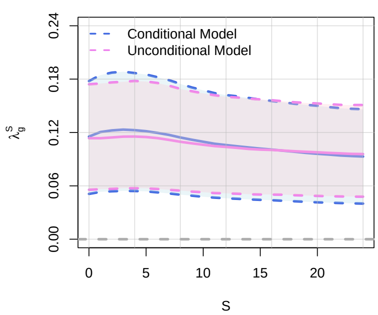
(p) Oil price change



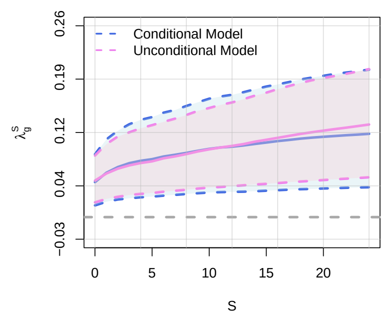
(q) TED spread



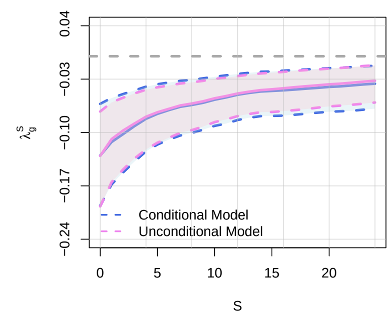
(r) HKM_{ntr}



(s) HKM_{tr}



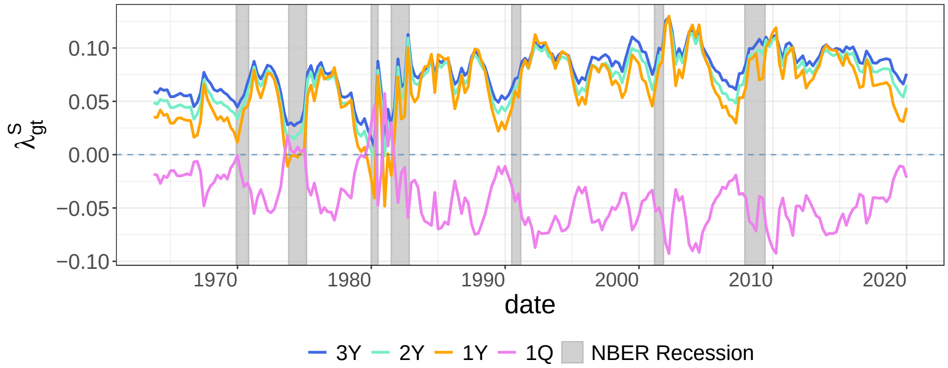
(t) PS liquidity



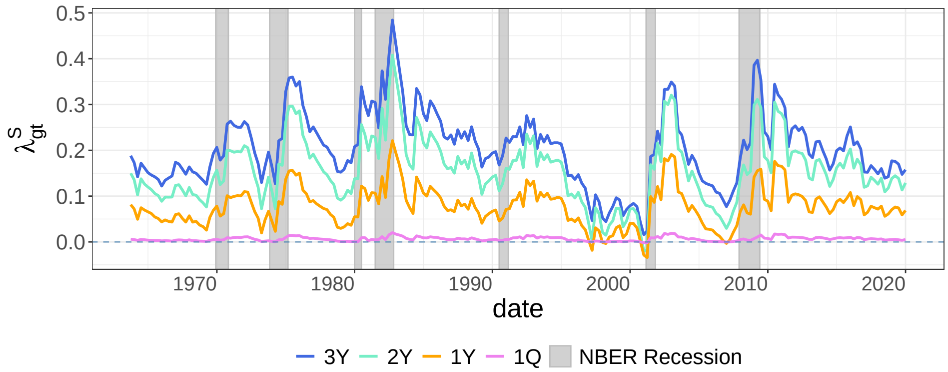
(u) $\Delta \log(\text{VIX})$

Figure IA.8: Term structure of unconditional risk premia in time-varying models with both lagged latent factors and external predictors: PE ratio of S&P 500, Term spread, default spread, and value spread

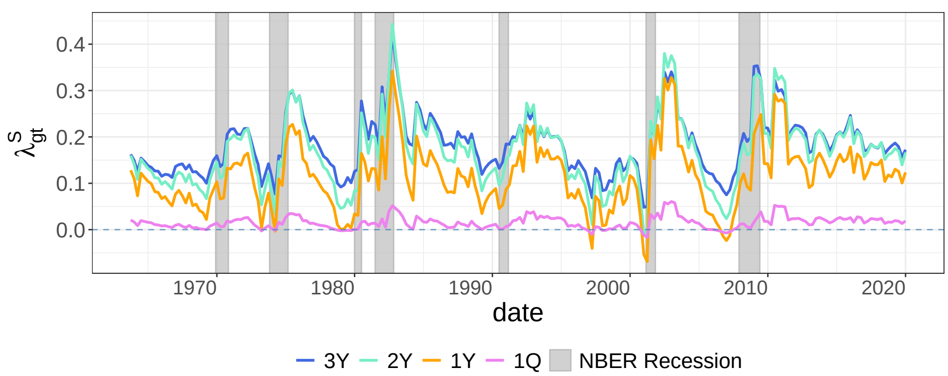
The figure plots the term structure of unconditional risk premia estimates using Propositions 1 and A2. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider the five-factor models for asset returns. For quarterly (monthly) factors, risk premia are estimated using a lag of 12 quarters (24 months) in g_t's equations. The blue dotted lines and the light blue shaded areas present the risk premia estimates and their 90% posterior credible intervals under the conditional models using the method described in Section 2.2. In the time-varying models, we model the dynamics of latent systematic factors as a VAR(1) process, with the PE ratio of the S&P500 index, term spread, default spread, and value spread as the external predictor. For comparison, we include the purple dotted lines and the related shaded areas, showing the risk premia estimated using the unconditional models described in Section 2.1. Definition and data sources of factors and test assets can be found in Appendix IA.3.



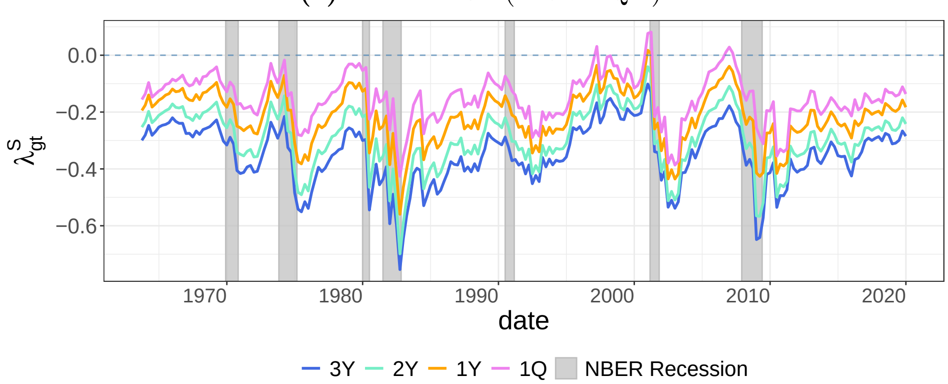
(a) Durable Consumption Growth



(b) Dividend growth of S&P500



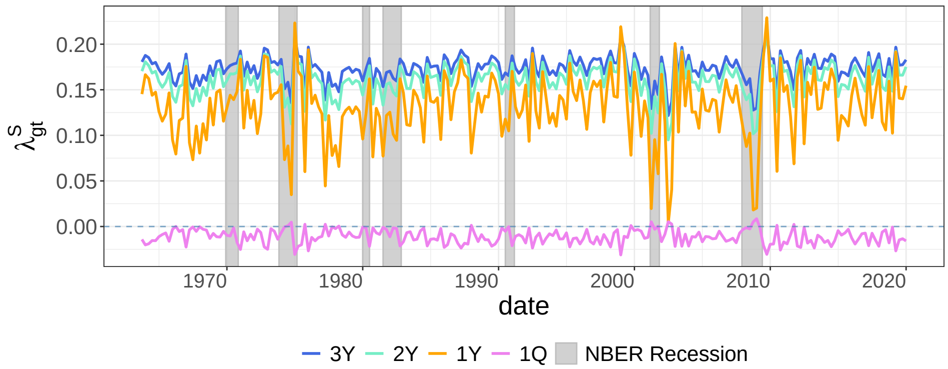
(c) Macro PC1 (FRED-QD)



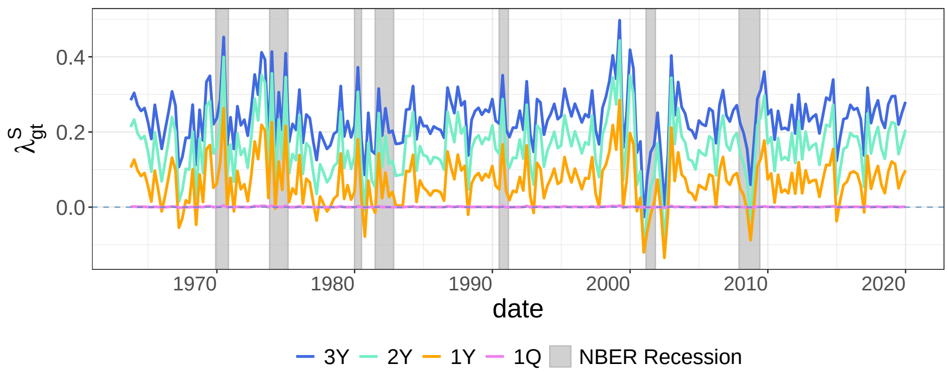
(d) Macro PC4 (FRED-QD)

Figure IA.9: Time-varying term structure of macroeconomic factor's risk premia

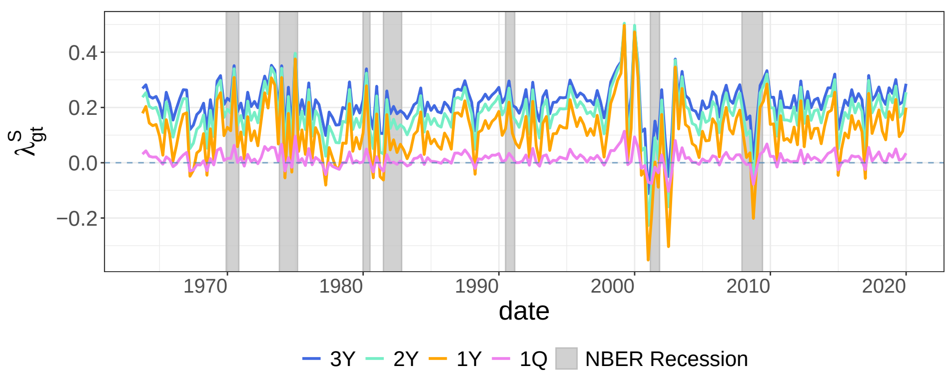
This figure plots the time-varying term structure of risk premia following the method in Section 2.2. Risk premia of latent factors are linear in four external predictors: PE ratio of S&P 500, Term spread, default spread, and value spread. Estimates are based on the composite cross-section of 275 Fama-French characteristic-sorted portfolios. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.



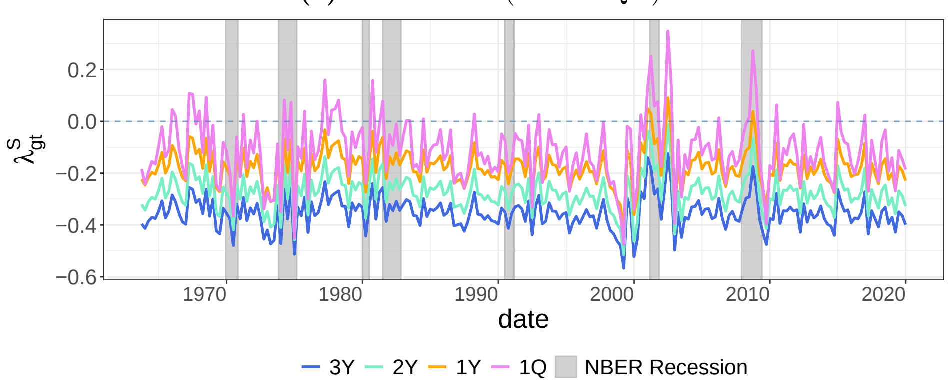
(a) Durable Consumption Growth



(b) Dividend growth of S&P500

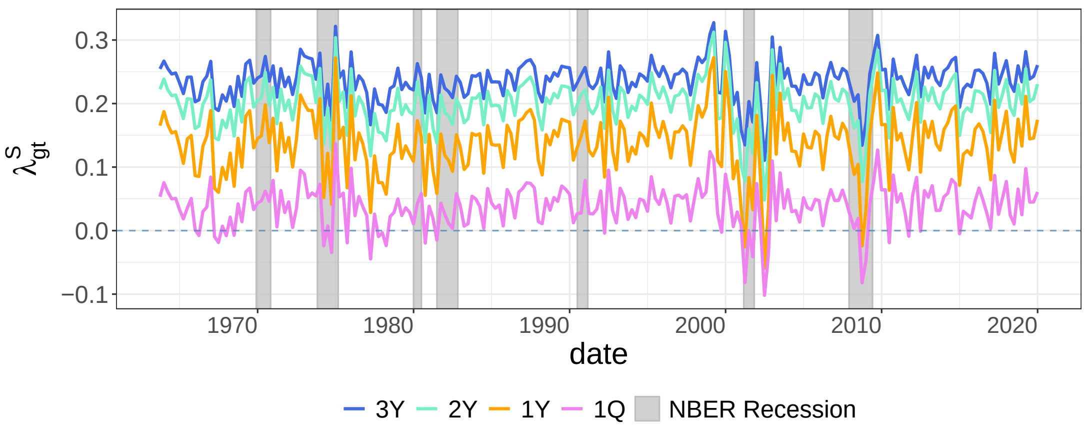


(c) Macro PC1 (FRED-QD)

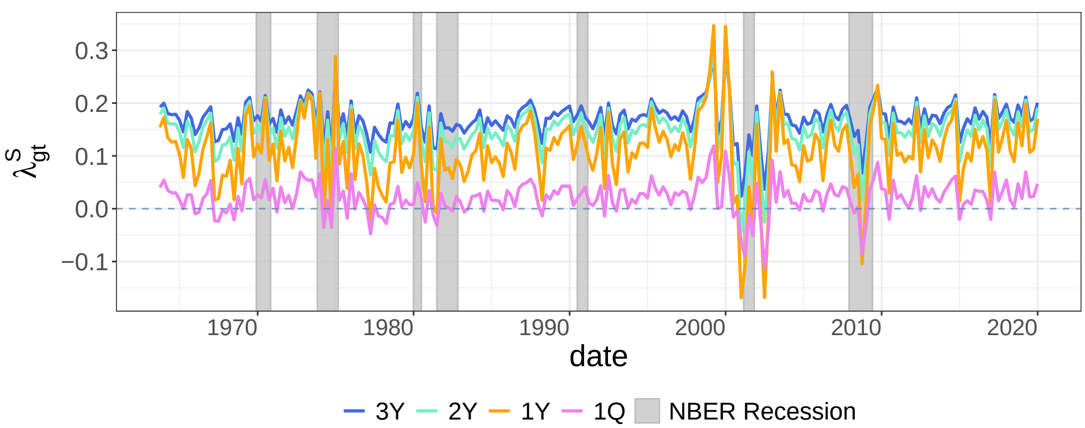


(d) Macro PC4 (FRED-QD)

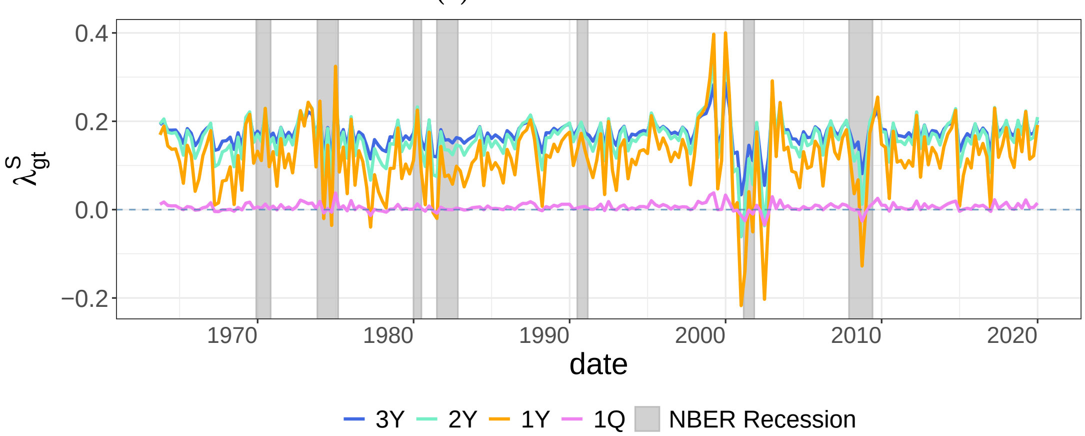
Figure IA.10: Time-varying term structure of macroeconomic factor's risk premia: VAR(1) model with only latent factors



(e) Nondurable Consumption Growth



(f) GDP Growth



(g) Industrial Production Growth

Figure IA.10: Time-varying term structure of macroeconomic factor's risk premia: VAR(1) model with only latent factors

This figure plots the time-varying term structure of risk premia following the method in Section 2.2 and a VAR(1) for the latent systematic risk factors. Estimates are based on the composite cross-section of 275 Fama-French characteristic-sorted portfolios. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

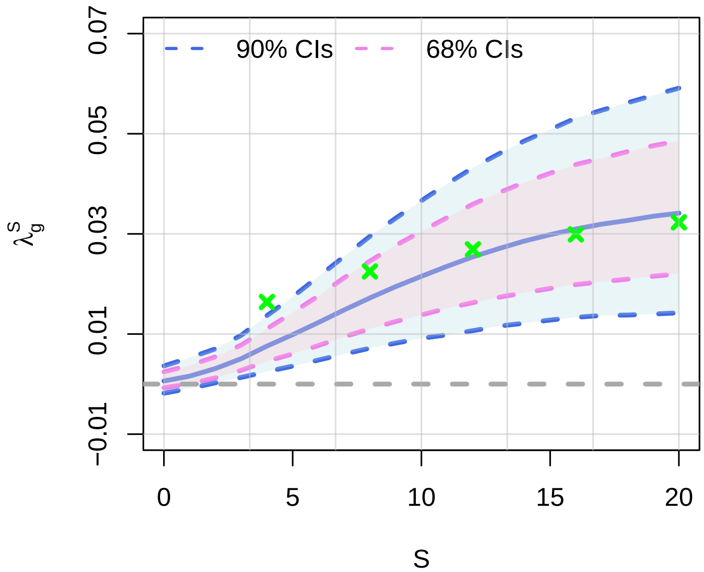


Figure IA.11: Term structure of unconditional dividend risk premia

This figure plots the term structure of dividend risk premia estimates. Unlike the estimates in Table 1, we do not standardize the dividend growth. The cross-section of test assets consists of 275 Fama-French characteristic-sorted portfolios. We consider a five-factor model for asset returns. We study the quarterly dividend growth, whose risk premia are estimated using a lag of 20 quarters in g_t's equations. In addition to the point estimates, we show the 68% and 90% Bayesian credible intervals, highlighted in pink and blue, respectively. The green crosses are the risk premia estimates obtained from Table 4 of Bansal et al. (2021).

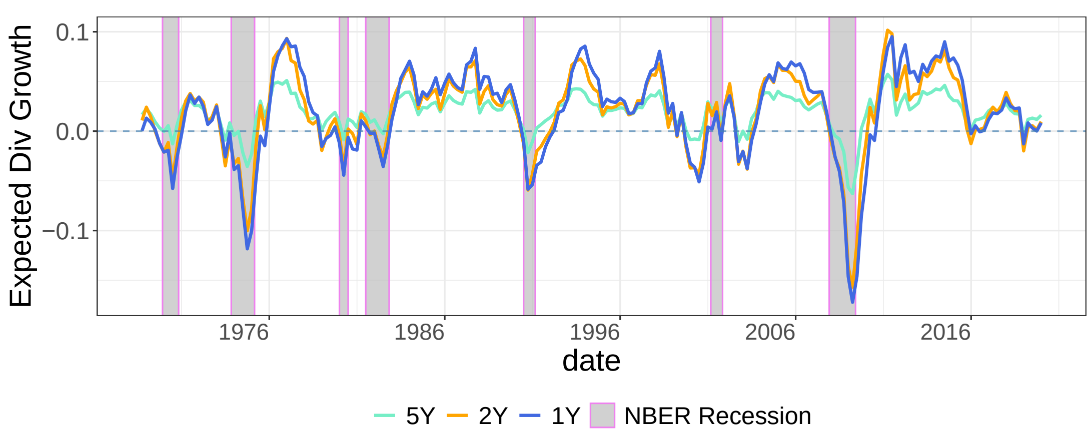


Figure IA.12: Expected dividend growth implied by the MA model: Constant risk premia model

This figure plots the time-varying expected dividend growth. The conditional mean of dividend growth is based on the MA model in equation (26), with $\bar{S} = 20$ quarters. Dividend risk premia are assumed to be constant. Estimates are based on the composite cross-section of 275 Fama-French characteristic-sorted portfolios. Definition and data sources of factors and test assets can be found in Internet Appendix IA.3.

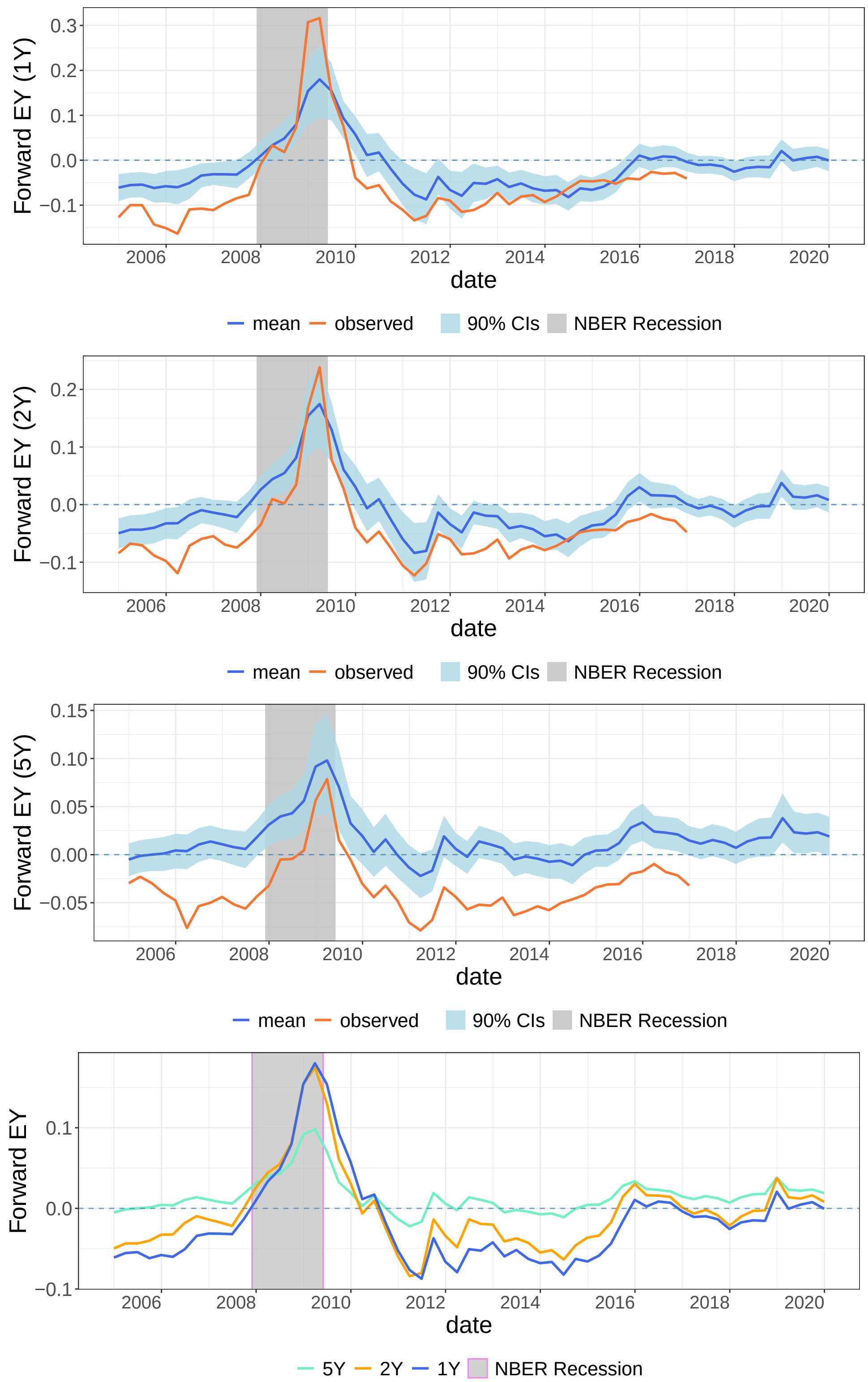


Figure IA.13: Time series of estimated forward equity yields: Constant dividend risk premia

This figure displays the time series of estimated forward equity yields based on our MA model with 20 lags. Dividend risk premia are assumed to be constant. We estimate a five-factor model of FF275 using the full sample from 1963Q3 to 2019Q4. We plot the estimates in the subsample from 2004Q4, the date from which we have the observed data of forward equity yields. The data of realised forward equity yields are from Bansal et al. (2021).