# Macro Strikes Back: Term Structure of Risk Premia*

Svetlana Bryzgalova† Jiantao Huang‡ Christian Julliard§

December 2025

**Abstract**

We provide a novel *priced* Wold representation theorem that sharply identifies shocks common to financial markets and the macroeconomy, their propagation, and the term structure of macro risk premia. Our identification leverages the fact that, in equilibrium models, asset prices are jump variables that reveal priced risks in large cross-sections of returns ($N \to \infty$). We find that macro factors' risk premia are strongly time-varying and countercyclical, with sharply increasing unconditional term structures. This pattern is driven by the slow propagation of a priced common shock, distinct from TFP, that captures most of the persistence in macroeconomic variables. Macro risk premia are negligible in the short run, yet grow to match equity risk premia at the business cycle horizon. This finding is not a mechanical byproduct of persistence: while GDP, consumption, industrial production, and employment exhibit upward-sloping term structures, other similarly persistent factors—e.g., VIX or intermediary-based variables—display downward-sloping or flat ones.

*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 Hengjie Ai, Nina Boyarchenko, Anna Cieslak, Mikhail Chernov, Richard Crump, Magnus Dahlquist, Marco Del Negro, Ian Dew-Becker, Greg Duffee, Xiang Fang, Marcelo Fernandes, Domenico Giannone, Bernardo Guimaraes, Ken Judd, Serhiy Kozak, Gabriele La Spada, Erik Loualiche, Ian Martin, Anna Mikusheva, Jun Pan, Lasse Pedersen, Elisabeth Proehl, Walt Pohl, Mirela Sandulescu, Elizaveta Sizova, Andrea Tamoni, Chenjie Xu, Paolo Zaffaroni, and seminar and conference participants at CEMFI, Copenhagen Business School, CUHK, CUHK Shenzen, DAFI Shanghai, Dartmouth Tuck, FGV São Paulo, FinEML, London Business School, London School of Economics, Maastricht University, Northwestern Kellogg, NY Fed, Renmin University of China, UT Sydney, UNSW, University of Lausanne, University of Sydney, Duke/UNC 2024 Asset Pricing conference, SoFiE 2024 conference, CICF 2024, NBER Summer Institute 2024, EEA-ESEM 2024, ESSFM Gerzensee 2024, 11th SAFE Asset Pricing Workshop, Finance Down Under 2025, FIRS 2025, SAIF annual conference 2025, WFA 2025, SITE 2025 and EFA 2025.

†Department of Finance, London Business School, and CEPR; 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.

*The general hypothesis [...] is that information about the production of a given period is spread across preceding periods and so affects the stock returns of preceding periods.*

— E. F. Fama (1990, p. 1096)

Macroeconomic risk, often associated with technology, consumption, or intermediary capital, is at the heart of most equilibrium-based asset pricing models. Yet, its reliable detection remains elusive: *i*) different time horizons often provide drastically different estimates of macro risk premia, *ii*) most empirical models are widely known to be misspecified and weakly identified at best, causing a fundamental inference problem, and *iii*) popular theoretical models often crucially rely on assumptions that are difficult, if not impossible, to test. All of these issues contribute to the empirical macro-finance disconnect.

We develop a new representation result: a *priced* Wold's decomposition theorem. We show that it identifies not only the joint comovement between nontradable macroeconomic factors and asset returns but also its propagation mechanism. This representation yields the whole (time-varying) term structure of risk premia, and sharp and internally consistent horizon-specific predictions for equilibrium macro-finance models.

Empirically, the links we uncover are much stronger and pervasive than previously thought: In a nutshell, we reconnect macro to finance. We find that most of the persistence in macroeconomic variables (e.g., industrial production, investment, consumption, hours worked, employment, and GDP growth) is driven by priced shocks. This generates increasing unconditional term structures with large and significant risk premia for macroeconomic factors. Furthermore, conditional macro risk premia are strongly time-varying and countercyclical. At business cycle frequencies (two–three years), macro risk is compensated by Sharpe ratios as large as that of the market portfolio (0.43–0.57 annualized). Remarkably, these risk premia are generated by the *very same* shocks that seem unpriced in the short run.

Our method provides sharp identification and reliable inference. Fundamentally, it is rooted in economic theory: Equilibrium asset prices are jump variables. Hence, news about current and *future* priced states are immediately reflected in prices, albeit they might manifest in the real economy with delay or over multiple periods.[1] Crucially, we turn the canonical approach of identifying priced macroeconomic risks on its head: Instead of attempting to price asset returns with (mismeasured and persistent) macroeconomic factors, we use a large cross-section of returns to identify the priced shocks and quantify their presence within the macro processes.

By uncovering the propagation of priced shocks in macro variables, we also explain why many risk factors seem to be statistically weak at quarterly frequency yet become strongly identified at longer horizons (e.g., Jagannathan and Wang (2007), Cohen et al. (2009)). Consider GDP growth, for example. Although contemporaneous asset return shocks account for only 4% of the variation in quarterly GDP growth, they contribute to over 25% of its time series variation at business cycle frequency. That is, contrary to the standard one-period inference, we find that a large part of the factor's conditional mean is driven by priced shocks propagating through macro time series (Adrian et al., 2019). In such a case, the term structure of risk premia effectively boosts the signal-to-noise ratio of priced shocks in nontradable factors: It sharply identifies the common priced component, which would be only weakly identified by conventional approaches.

We find that most macro variables are driven by (almost) identical priced shocks—as implied by our novel theoretical Wold decomposition and tested in the data. These shocks: *i*) explain a large part of time series variations in macro variables (20%-47%) and most of their predictability, *ii*) imply a correlation between their priced components in the 45-95% range, *iii*) command almost identical annualized Sharpe ratios (0.43-0.57), *iv*) generate extremely similar macro response to financial market shocks at business cycle frequencies, and *v*) produce a clear business cycle pattern in the conditional means of macro factors. That is, financial markets are first-order important for macro dynamics (Jordà et al., 2017), and macro news is of first-order relevance for financial markets (Chen et al., 1986). In many models, such link would be driven by productivity shocks. Yet, surprisingly, as in Angeletos et al. (2020), we find no evidence of canonical measures of TFP driving this commonality, as TFP shocks do not seem to be priced.

Our priced Wold representation requires only the covariance stationarity of the pricing kernel and leverages the fact that many nontradable factors are persistent. Our identification relies on three key ingredients: 1) a novel 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 (Chamberlain and Rothschild, 1983), and 3) the hierarchical Bayesian inference method of Bryzgalova et al. (2023), which recovers both time series and cross-sectional properties of risk factors. This approach boosts the signal-to-noise ratio of macro quantities by leveraging their term structure, efficiently separates priced from unpriced components of macro variables, and nests *all* popular macro-finance models.

[1] Slow propagation of priced shocks into real quantities has been motivated in the literature as being caused by measurement errors and/or real or informational friction, see, e.g., Kydland and Prescott (1982), Breeden et al. (1989), Fama (1990), Wilcox (1992), Bansal and Yaron (2004), Hansen et al. (2008), Abel et al. (2013), Collin-Dufresne et al. (2016).

Echoing the maximum-share identification strategy of the VAR literature (Faust, 1998; Uhlig, 2003; Barsky and Sims, 2011; Francis et al., 2014; Angeletos et al., 2020), we extract the best linear combination of priced financial shocks that explains the maximal fraction of time-series variation in macro variables. However, instead of fragile autoregressive structures, we use a flexible MA-based representation that allows us to robustly identify the propagation of priced shocks into macro quantities. Our approach, like the local projection framework (Jordà, 2005), robustly recovers the impulse response functions of macro factors to shocks spanned by financial markets *without* postulating a stringent (V)AR structure (Olea et al., 2024).

Unlike traditional macro-finance papers, we leverage a very large cross-section of asset returns to separate priced and unpriced innovations and restore reliable inference on risk premia. Similar to Giglio and Xiu (2021), we use 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 indeterminacy and can, therefore, be uniquely recovered from the data. Our estimator is robust to omitted variable bias, measurement error, and weak identification. Extensive simulations confirm that the Bayesian credible intervals provide proper posterior coverage.[2] Moreover, we show that our MA-based approach is key for identifying the priced component in factors and boosts statistical power.

Our findings are not a simple byproduct of factor persistence. Similarly persistent factors can have increasing (liquidity of Pástor and Stambaugh (2003)), flat (intermediary factor of He et al. (2017)), or decreasing (VIX) term structures, or no significant risk premia at all (capital share growth of Lettau et al. (2019)). We map risk premia over different horizons 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, e.g., industrial production, GDP, and unemployment, are as large as that of the market. Importantly, our results do not rely on ad-hoc frequency decompositions, but are instead fully determined by the identified propagation of priced shocks through macro variables, reconciling the seemingly discordant findings across frequencies in the previous literature.

We also investigate the relationship between the term structure of risk premia and forward equity yields implied by dividend strip data. Since dividend strips are virtually the only equity cash flows for which a term structure of prices is directly observable, these data have attracted

[2] Methodologically, our approach is closely linked to the Bayesian empirical macro literature, e.g., Primiceri (2005), Crump et al. (2021), Del Negro and Primiceri (2015).

considerable attention and generated its own literature (e.g., van Binsbergen et al. (2012)). Ex-ante, our model is not guaranteed to generate equity yields matching observed data (Bansal et al., 2021). Strikingly, ex-post, our countercyclical estimates of forward equity yields are very close to those implied by the strips data—a strong validation of our approach.

Our method jointly models macro factors and return dynamics over different horizons and provides coherent insights into the whole term structure of risk premia. Tackling an inference problem in this setup is challenging, if not infeasible, in frequentist estimation. Instead, we develop a 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.

We uncover two first-order methodological fallacies responsible for blurring the link between asset returns and the real economy. First, it is common practice to build mimicking portfolios for macro factors and measure risk premia as their average excess returns. This typically results in very low risk compensations at monthly/quarterly frequency, leading researchers to conclude that macro risk is either not priced or weakly identified (Herskovic et al., 2019). We show, theoretically and empirically, that such mimicking portfolios cannot recover the term structure of risk premia of most macro variables. This happens because the projection of nontradable factors on the space of returns omits the slow propagation of priced shocks—which are crucial to characterize macro risk premia. This implies that for persistent factors (such as GDP and unemployment), canonical mimicking portfolios mismeasure their risk premia. The data confirm this pitfall. Instead, our method leverages the predictability of macro quantities by financial markets (Liew and Vassalou, 2000) and constructs portfolios that capture the forward-looking priced component of macro variables (Lamont, 2001).

Second, the literature has long recognized the persistent nature of many nontradable factors. Hence, researchers often extract their AR(1) innovations and measure their risk premia.[3] However, this common procedure fails to correctly recover risk premia because the overall factor persistency can also be caused by non-priced shocks (e.g., measurement error). Furthermore, the priced component of macro variables is often different from an AR(1). Thus, relying on AR(1) residuals leads to biased estimates. Instead, our method effectively separates priced and unpriced dynamics in macro variables without ad hoc assumptions. Our results have important implications for dynamic factor models and VARs in empirical macro-finance, which typically

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

do not separate priced and unpriced drivers of conditional means (Stock and Watson (2011)).

The remainder of the paper is organized as follows. Next, we review the most closely related literature. Section 1 outlines our method and its properties, while Section 2 provides simulation evidence in realistically small samples. Section 3 presents our empirical findings, and Section 4 concludes. Additional details, results, and proofs are reported in the Internet Appendix.

### (Additional) Closely Related Literature

Equilibrium macro-finance models have sharp and salient predictions for the term structures of risk premia of macro factors. For instance, the habit model of Campbell and Cochrane (1999) predicts flat term structures of risk premia for consumption and dividend growth (see Figure A1). Yet these same factors command upward-sloping term structures in the long-run risk model of Bansal and Yaron (2004). These predictions differ from each other because they rely on different ad hoc assumptions on cash flow dynamics and investors' preferences (Zviadadze, 2021). Instead, our framework allows us to robustly elicit the term structure of risk premia directly from the data in a flexible model-free way (Backus et al., 2018; Giglio et al., 2023).

Methodologically, our paper is close to the Bayesian empirical macro literature that estimates time-varying dynamics and structural shocks in VARs.[4] Like Primiceri (2005), Del Negro and Primiceri (2015), and Crump et al. (2021), we work in a high-dimensional setting and use a Gibbs sampler to obtain the joint posterior of latent states, factor loadings, and shock variances. This approach yields credible sets for all key objects of interest that are straightforward to compute and interpret. However, our method differs in two important ways. First, instead of specifying a VAR for a small set of macro aggregates, we build on a novel priced Wold decomposition that yields a flexible MA representation. That is, we work directly in the space of impulse-response functions. Consequently, our shock identification strategy relies on cross-sectional pricing restrictions implied by the stochastic discount factor (rather than on exclusion or sign restrictions as in the typical VAR). Second, most of the existing literature uses Bayesian VARs mainly to forecast and trace the transmission of monetary and macro shocks (Del Negro and Schorfheide, 2004; Sims and Zha, 2006; Bańbura et al., 2010). Our approach, instead, allows to separate priced and unpriced components of macro series. This, in turn, recovers in one coherent system the time-varying term structure of macro risk premia across horizons, and thus provides the macro-finance counterpart to the Bayesian VAR tradition.

[4] Sims (1992), Cogley and Sargent (2005), Smets and Wouters (2007), Giannone et al. (2015).

Our paper naturally relates to the large literature, starting as far back as Black et al. (1972), Shanken (1992), and Fama and French (1992), on the estimation and inference of risk premia in linear factor models. These approaches are typically fragile to model misspecification (Kan et al. (2013)). As a result, recent literature formally tackles the omitted variable problem by using systematic risk factors identified in large cross-sections of asset returns (typically via PCA-based methods, e.g., Chamberlain and Rothschild (1983), Connor and Korajczyk (1986, 1988), Kozak et al. (2020)). Our method also explicitly controls for omitted variable bias, but unlike the existing robust approaches, we leverage the Bayesian hierarchical framework of Bryzgalova et al. (2023) to additionally incorporate the dynamics of factors and returns. This allows us to elicit the term structure of risk premia in an internally consistent manner and increase statistical power. Note also that our use of a combination of flat and conjugate priors recovers the frequentist estimates as posterior modes, making our results directly comparable to the misspecification-robust frequentist estimators, e.g., Giglio and Xiu (2021).

The empirical macro-finance literature has struggled for a long time to rationalize the drastically different estimates for macroeconomic risk premia obtained at different frequencies (see, e.g., the disagreement among estimates of consumption risk premia, Mehra and Prescott (1985), Lettau and Ludvigson (2001), Jagannathan and Wang (2007), Ortu et al. (2013), Kleibergen and Zhan (2020), and Bandi and Tamoni (2023)).[5] We show that these seemingly discordant findings are generated by the multi-period, horizon-specific, response of macro variables to priced financial shocks.

Unlike Parker and Julliard (2005), who show that using long-horizon consumption responses improves cross-sectional pricing and lowers risk aversion estimates in a CRRA model, we not only document horizon-specific pricing of macroeconomic risk, but crucially we identify its economic driver and propagation mechanism. This allows us to uncover a fallacy in the common narrative attributing frequency-based pricing to horizon-specific attitudes toward risk (e.g., Andries et al. (2024)): we show instead that the term-structure of risk premia is generated by the *same* shocks slowly propagating within the macroeconomy. Hence, as we find in the data, macro risks can have increasing (e.g, GDP), flat (e.g., intermediary capital as measured in He et al. (2017)), or even decreasing (VIX) term structures of risk premia, as their shape is not

[5]Similarly, risk premia estimates for the VIX index have been found to vary with the horizon, see, e.g., Eraker and Wu (2014), Dew-Becker et al. (2017), and Johnson (2017)). Furthermore, Chernov et al. (2021) documents striking horizon-specific mispricing even for linear models with only tradable factors (e.g., the Fama-French five-factor model).

determined by their overall level of persistence, but rather by how the *priced* shocks propagate within their processes.

This key distinction between priced and unpriced shocks is exactly what our novel Wold representation yields. As Bryzgalova et al. (2025), we model macro quantities as flexible MA representations, but unlike them we leverage the asset pricing restrictions to isolate the role of shocks that command significant risk premia in the cross-section of asset returns. That is, we turn the identification problem on its head: instead of recovering the latent drivers of macro quantities (as in Bryzgalova et al. (2025) and Stock and Watson (2011)) and exploring their pricing and predictive ability, we use a large cross-section of returns to identify the (tradable) SDF, and pinpoint the propagation of its shocks in macro quantities. That is, we identify the shocks in financial markets using higher frequency data, and trace the response of macro quantities to these shocks (see, e.g., Kuttner (2001), Bernanke and Kuttner (2005), Gomez Cram et al. (2025), and Bianchi et al. (2022)). This approach also circumvents the weak factor problem in asset pricing that blurs the link between finance and macroeconomic quantities (Kan and Zhang (1999a,b) and Kleibergen and Zhan (2020)).

# 1 Theory and Method

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, investment, 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 1.2. 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. 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 Jensen's inequality.[6] In addition, we assume that each asset's pricing error, $\alpha_i$, 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 the latent factors $\tilde{\boldsymbol{v}}_t$:[7]

$$m_t - \kappa_m = -\boldsymbol{\lambda}_{\tilde{v}}^\top \tilde{\boldsymbol{v}}_t =: \varepsilon_t^m \tag{3}$$

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

To model each covariance-stationary factor $g_t$ as the sum of two components—a priced and an unpriced component—we introduce the following novel representation result.

**Theorem 1** (Priced Wold Decomposition)**.** *Let the demeaned (log) stochastic discount factor $\{\tilde{m}_t\}_{t\in\mathbb{Z}}$ and the scalar process $\{g_t\}_{t\in\mathbb{Z}}$ with mean $\mu_g$ be covariance-stationary with finite second moments on $(\Omega, \mathcal{F}, P)$. Define $\mathcal{M}_t := \overline{\text{span}}\{\tilde{m}_s : s \leq t\} \subset L^2(\Omega, \mathcal{F}, P)$ and $\varepsilon_t^m := \tilde{m}_t - \Pi_{\mathcal{M}_{t-1}}(\tilde{m}_t)$ where $\Pi_{\mathcal{S}}(\cdot)$ denotes the orthogonal projection onto a closed subspace $\mathcal{S} \subset L^2(\Omega, \mathcal{F}, P)$, and let $\sigma_\varepsilon^2 := \text{Var}(\varepsilon_t^m) > 0$. Assume that for all $t$, the mean-zero component of $g_t$ can be written as $g_t - \mu_g = \tilde{g}_t + \tilde{w}_t^g$, $\tilde{g}_t \in \mathcal{M}_t$, $\tilde{w}_t^g \perp \mathcal{M}_t$. Then there exists a sequence $\{\theta_j\}_{j=0}^\infty$ satisfying $\sum_{j=0}^\infty \theta_j^2 < \infty$, and a process $w_t^g$ orthogonal to $\tilde{m}_t$, such that*

$$g_t = \mu_g + \sum_{j=0}^{\infty} \theta_j\, \varepsilon_{t-j}^m + w_t^g, \qquad \theta_j = \frac{\mathbb{E}[(g_t - \mu_g)\, \varepsilon_{t-j}^m]}{\sigma_\varepsilon^2}, \tag{4}$$

*and this representation is unique.*

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

[7] Our paper does not model the risk-free rate and hence the dynamics of conditional mean of the SDF. For convenience, we normalize the mean of $m_t$ to be an unknown constant $\kappa_m$.

We prove Theorem 1 in Appendix A.1, but its core idea is intuitive. If we decompose the variable $g_t$ into a purely priced component—containing no idiosyncratic information orthogonal to systematic risk—and an unpriced component, then by definition the former must lie in the linear span of the SDF (Hansen and Richard, 1987). Consequently, by virtue of the Wold theorem applied to the SDF, we can express the priced part of $g_t$ as a moving average of current and lagged SDF innovations.

Based on the representation result in equation (4), and using the SDF in equation (3), we model the factor $g_t$ as the sum of a moving average of asset-return shocks and additional shocks (such as measurement error) not spanned by financial markets:

$$g_t = \mu_g + \sum_{s=0}^{\bar{S}} \tilde{\rho}_s \underbrace{\tilde{\boldsymbol{\eta}}_g^\top \tilde{\boldsymbol{v}}_{t-s}}_{f_{t-s}} + w_{gt}, \quad \tilde{\boldsymbol{\eta}}_g^\top \tilde{\boldsymbol{\eta}}_g = 1, \tag{5}$$

where $\mu_g$ is the unconditional mean, $\tilde{\boldsymbol{\eta}}_g \propto \boldsymbol{\lambda}_{\tilde{v}}$ (a restriction we explore in Sections 3.1.1 and 3.1.2), $f_t$ is the spanned component that may drive both $g_t$ and asset returns, $\{\tilde{\rho}_s\}_{s=0}^{\bar{S}}$ is a square-summable sequence of scalars, 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, Theorem 1 ensures existence of the above representation, possibly with $\bar{S} = \infty$. Using a finite $\bar{S}$ introduces only a finite approximation error relative to the true priced component, due to the square-summability of the MA coefficients in Theorem 1.

Since $f_t$ is a white noise innovation, we can also interpret $\{\tilde{\rho}_s\}_{s=0}^{\bar{S}}$ as $g_t$'s impulse responses to the asset return shock $f_t$.[8] That is, $\tilde{\rho}_s = \mathbb{E}\left[g_{t+s}|f_t = 1; \{f_{t-j}\}_{j=1}^{\infty}\right] - \mathbb{E}\left[g_{t+s}|f_t = 0; \{f_{t-j}\}_{j=1}^{\infty}\right]$, and $\tilde{\rho}_s$ 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., 2025). Furthermore, since we make use of the existence of a MA representation in the construction of $\mathbb{E}\left[g_{t+s}|f_t = 0; \{f_{t-j}\}_{j=1}^{\infty}\right]$, we learn from the data about $\tilde{\rho}_s$ not only from $g_{t+s}$ but also from all other leads and lags of $g_t$, arguably gaining efficiency. Moreover, our approach directly accounts for the MA structure implied by a sequence of linear projections[9] and uses only one estimating equation (equation (5), instead of $\bar{S}$ linear regressions

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

[9] Lusompa (2023) shows that the autocorrelation process of LP can be written as a Moving Average process

with correlated residuals), hence greatly simplifying inference.

Several features of equation (5) 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) and Lamont (2001)). Third, when $g_t$ 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).[10] Fourth, as shown in Internet Appendix IA.2, model selection of the type of (low-dimensional) ARMA processes postulated in the macro-finance literature (e.g., consumption, TFP, and profitability) is fragile at best and rarely recovers the true process. Therefore, it is more robust to work with the general long MA representation. As shown in Bryzgalova et al. (2025) for the consumption process, the conditional mean dynamics and, hence, the innovations needed to derive the pricing implications, can be accurately captured using a long MA representation as in equation (5).

Our representation result is very general and covers all macro-finance models (as long as the SDF is covariance stationary).[11] We now use several examples to illustrate how the framework in Equations (1)–(5) maps into canonical models imposing particular parametric restrictions.[12]

**Example 1.** *Adrian et al. (2014) measure a financial intermediary SDF, i.e., $m_t = \kappa_m - \lambda \cdot LevFac_t$, where $LevFac_t$ is the shock to the leverage of security broker-dealers. Our framework maps into theirs with: $\tilde{v}_t = f_t = LevFac_t$, $\bar{S} = 0$, $\tilde{\rho}_0 = 1$, and $g_t$ is a noisy proxy for $LevFac_t$.*

**Example 2.** *In the canonical long-run risk model of Bansal and Yaron (2004), the log consumption growth is modeled as $g_t = \Delta c_t = x_{t-1} + \sigma_{t-1}\eta_t$, where $\sigma_{t-1}$ is the stochastic volatility process, $x_{t-1}$ is the conditional consumption mean following an AR(1) process, $x_t = \rho_x x_{t-1} + \varphi_e \sigma_{t-1} e_t =$*

of the Wold errors and impulse responses, and accounting for this dependency leads to more efficient estimates.

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

[11]Our formulation allows shocks to be driven by a jump process. Nevertheless, in the empirical applications, we model all shocks as driven by continuous distributions since, at the quarterly and monthly frequencies we focus on, would imply modelling quite small jumps that are well approximated by, and are hard to distinguish from, a continuous process (see, e.g., Aït-Sahalia (2004)).

[12]Additional examples, e.g., Croce (2014) and Belo and Li (2023), are discussed in Internet Appendix IA.2.

$\sum_{s=0}^{\infty} \varphi_e \rho_x^s \sigma_{t-s-1} e_{t-s}$, *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.,* $\varepsilon_t^m := m_t - \mathbb{E}_{t-1}(m_t) = \lambda_{m,\eta}\sigma_{t-1}\eta_t - \lambda_{m,e}\sigma_{t-1}e_t - \lambda_{m,\omega}\sigma_\omega\omega_t$ *(*$\sigma_\omega\omega_t$ *is the shock to* $\sigma_t^2$*, see Internet Appendix IA.5 for details). The SDF in equation (3) maps into the Bansal and Yaron (2004) model by 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)* $\boldsymbol{\lambda}_{\tilde{v}} = (\lambda_{m,\eta}, -\lambda_{m,e}, -\lambda_{m,\omega})^\top$*. Furthermore, the fundamental priced Wold representation*[13] *in equation (4) yields the coefficient restrictions:* $\bar{S} = \infty$*,* $\tilde{\rho}_0 = \frac{\lambda_{m,\eta}\,\mathbb{E}[\sigma_{t-1}^2]}{\sigma_\varepsilon^2}$*,* $\tilde{\rho}_{j\geq 1} = \frac{-\rho_x^{j-1}\varphi_e\,\lambda_{m,e}\,\mathbb{E}[\sigma_{t-1}^2]}{\sigma_\varepsilon^2}$ *and* $\sigma_\varepsilon^2 = \left(\lambda_{m,\eta}^2 + \lambda_{m,e}^2\right)\mathbb{E}[\sigma_{t-1}^2] + \lambda_{m,\omega}^2\sigma_\omega^2$*.*

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, $\lambda_g = -\text{cov}(g_t, m_t)$.[14] When $g_t$ is a traded log excess return, the fundamental asset pricing equation, $\mathbb{E}\big[\exp(m_t + g_t + r_f)\big] = 1$, 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 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:

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

There are two ways to interpret equation (6). First, $\lambda_f$ is the risk premium of the spanned component ($f_t = \tilde{\boldsymbol{\eta}}_g^\top\tilde{\boldsymbol{v}}_t$) driving both 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 return shocks $f_{t-1\to t+S}$. Hence, $\lambda_g^S$, 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 $\lambda_g^S$ as the risk premium of the *horizon-specific* mimicking portfolio hedging against $g_{t-1\to t+S}$, with projection-based portfolio weights $\boldsymbol{w}^{MP} = \text{cov}(\boldsymbol{r}_{t-1\to t+S})^{-1}\text{cov}(\boldsymbol{r}_{t-1\to t+S}, g_{t-1\to t+S})$. The risk premium of this portfolio, normalized

[13]Recall that the Wold representation of a random variable is *not* algebraically identical to the random variable itself. Nevertheless, it is identical in the Hilbert space sense (i.e., it represents the same element of $L^2(\Omega, \mathcal{F}, P)$) as it preserves all first and second moments, hence yielding the exact same pricing implications and impulse responses of the original variable.

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

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 1.2).

Using Lemma A.2 in the Appendix, we can simplify the risk premium of $g_t$'s mimicking portfolio obtaining 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 $m_{t-1\to t+S} = \sum_{\tau=0}^{S} m_{t+\tau-1,t+\tau} = (1+S)\kappa_m - \boldsymbol{\lambda}_{\tilde{v}}^\top \tilde{\boldsymbol{v}}_{t-1\to t+S}$. 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 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 $m_t \propto -\lambda_{mkt} r_t^{mkt}$. For any factor $g_t$ the term structure of its risk premia is:*

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

*where the forward betas, $\beta_S^g \equiv \frac{\text{cov}(g_{t-1+s\to t+s}, r_t^{mkt})}{\sigma_{mkt}^2}$, capture the predictability of $g$. The term structure of risk premia is determined by how the same priced shock propagates through the factor. The mimicking portfolio based on the single-period market beta ($\beta_0^g$) is uninformative about the multi-period risk premia, since it ignores the information 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. This rotation invariance property (Giglio and Xiu (2021)) can be easily seen by rewriting the model as follows:

$$\boldsymbol{r}_t = \boldsymbol{\alpha} + \underbrace{\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\boldsymbol{\lambda}_{\tilde{v}}}_{\boldsymbol{\lambda}_v} - \frac{1}{2}\boldsymbol{\Upsilon}_r + \underbrace{\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\tilde{\boldsymbol{v}}_t}_{\boldsymbol{v}_t} + \boldsymbol{w}_{rt}, \quad g_t = \mu_g + \sum_{s=0}^{\bar{S}}\tilde{\rho}_s\underbrace{\tilde{\boldsymbol{\eta}}_g^\top\boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_g^\top}\underbrace{\boldsymbol{H}\tilde{\boldsymbol{v}}_{t-s}}_{\boldsymbol{v}_{t-s}} + w_{gt}, \text{ and}$$

$$m_t = \kappa_m - \boldsymbol{\lambda}_v^\top(\boldsymbol{H}^{-1})^\top\boldsymbol{H}^{-1}\boldsymbol{v}_t = \kappa_m - \boldsymbol{\lambda}_v^\top\boldsymbol{\Sigma}_v^{-1}\boldsymbol{v}_t, \quad \lambda_g^S = \frac{\sum_{\tau=0}^{S}\sum_{s=0}^{\tau}\tilde{\rho}_s}{1+S}\cdot\underbrace{\tilde{\boldsymbol{\eta}}_g^\top\boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_g}\underbrace{\boldsymbol{H}\boldsymbol{\lambda}_{\tilde{v}}}_{\boldsymbol{\lambda}_v}. \tag{7}$$

Therefore, the most important quantities, the impulse response coefficients $\{\tilde{\rho}_s\}$ and the risk premia $\lambda_g^S$, are point identified.

**Remark 1.** *$\lambda_f$ 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 $m_t = \kappa_m - \boldsymbol{\lambda}_v^\top \boldsymbol{\Sigma}_v^{-1} \boldsymbol{v}_t = \kappa_m - \lambda_f f_t - \boldsymbol{\lambda}_u^\top (\boldsymbol{H}_1^\top \boldsymbol{H}_1)^{-1} \boldsymbol{u}_t$, 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 $\mathrm{var}(m_t) = \lambda_f^2 + \mathrm{var}(\boldsymbol{\lambda}_u^\top (\boldsymbol{H}_1^\top \boldsymbol{H}_1)^{-1} \boldsymbol{u}_t)$. Hence, $\lambda_f$ can be interpreted as the (per period) model-implied Sharpe ratio of the $f_t$ shock; $\lambda_f^2/\mathrm{var}(m_t)$ quantifies the relative importance of $f_t$ in the SDF, where $f_t$ is constructed to have the largest power in the log SDF to explain the cross-section of average returns.*

Estimating the confidence bands—or better, the statistical uncertainty—of $\lambda_g^S$ is challenging in the frequentist framework. Specifically, $\lambda_g^S$ is a function of $\{\tilde{\rho}_s\}_{s=0}^{\bar{S}}$, $\boldsymbol{\eta}_g$, and $\boldsymbol{\lambda}_v$, where the first two sets of parameters depend on each other. Hence, the frequentist asymptotic covariance matrix of $\lambda_g^S$ is quite complex despite its closed-form expression outlined above. Consequently, we adopt a Bayesian framework to provide valid inference for all model parameters.

## 1.1 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:

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

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

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

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 $\boldsymbol{\mu}_r$.[15] According to equation (10), the error terms, $w_{gt}$ and $\boldsymbol{w}_{rt}$, are

[15] 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

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 (5). 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.

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 almost 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, $\lambda_v$, 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

$$m_t = \kappa_m - \boldsymbol{\lambda}_v^\top \boldsymbol{\Sigma}_v^{-1} \boldsymbol{v}_t \;\Rightarrow\; \tilde{\boldsymbol{\mu}}_r = \boldsymbol{\beta}_v \boldsymbol{\lambda}_v, \tag{12}$$

where $\tilde{\boldsymbol{\mu}}_r = \boldsymbol{\mu}_r + \frac{1}{2}\boldsymbol{\Upsilon}_r$, with both $\boldsymbol{\mu}_r$ and $\boldsymbol{\Upsilon}_r$ estimated in the time series step. Recall that we nevertheless allow for pricing errors as outlined in (2). With extensive simulation studies, we show in Section 2 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

disappear. Consequently, the credible intervals for $\boldsymbol{\mu}_r$ will be too tight if we do not directly model $\boldsymbol{\mu}_v$.

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).

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 $\boldsymbol{v}_t$ can be obtained by inverting their relationship with asset returns. Finally, conditional on parameters and latent factors in the time series layer, the distribution of 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.1.2.

**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 ($\sigma_{wg}^2$, $\boldsymbol{\rho}_g$, and $\boldsymbol{\eta}_g$) follow the normal-inverse-gamma distribution in equations (IA.1)–(IA.3) of Internet Appendix IA.1.2. 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.1.2.*

*(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*

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

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

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

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

*(4) Conditional on the 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 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.[16]

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 $N \to \infty$. 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 $g_t$. Furthermore, since our identification approach finds the best linear combination of financial shocks $\boldsymbol{v}_t$, $f_t = \boldsymbol{\eta}_g^\top \boldsymbol{v}_t$, such that $\{f_{t-s}\}_{s=0}^{\bar{S}}$ explain the maximal fraction of time-series variation in $g_t$, our method echoes the max-share identification strategy of Faust (1998), Uhlig (2003), Barsky and Sims (2011), Francis et al. (2014), Angeletos et al. (2020). Moreover, as per Proposition 1, theoretically $\tilde{\boldsymbol{\eta}}_g \propto \boldsymbol{\lambda}_{\tilde{v}}$. We exploit this testable restriction in the analysis in Section 3.1.2.

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

[16]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 $\{v_{t-s}\}_{s=0}^{\bar{S}}$.

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., $N \leq 50$), we can 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 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 GMM and Fama-MacBeth regressions suffer from weak identification (see, e.g., Kan and Zhang (1999a,b)), particularly for macro factors. One contribution of our method 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 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 macro factors.

## 1.2 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,[17]

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

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

where $\mathbb{E}_t$ denotes the time $t$ conditional expectation, and $\mathrm{var}_t(r_{i,t+1})$ is the conditional variance. We consider homoskedastic asset returns, hence $\mathrm{var}_t(r_{i,t+1})$ is constant over time. We define $\boldsymbol{\Upsilon}_r$ as $\big(\mathrm{var}_t(r_{1,t+1}), \ldots, \mathrm{var}_t(r_{N,t+1})\big)^\top$. Leveraging Hansen and Jagannathan (1991), we focus on the conditional SDF projections on the space of returns as follows:

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

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 $\boldsymbol{\mu}_{\tilde{v},t-1} \equiv \mathbb{E}_{t-1}\left[\tilde{\boldsymbol{v}}_t\right]$; hence $\boldsymbol{\mu}_{\tilde{v},t-1} \perp \boldsymbol{\epsilon}_{\tilde{v}t}$. We normalize the innovations to the latent factors such that $\mathrm{cov}(\boldsymbol{\epsilon}_{\tilde{v}t}) = \boldsymbol{I}_K$.

As previously, unconditional mean returns are partially explained by $\boldsymbol{\beta}_{\tilde{v}}$ in equation (2). The only additional required assumption is that the eigenvalues of $\mathrm{cov}(\boldsymbol{\mu}_{\tilde{v},t-1})$ are bounded. This yields an SDF where $\boldsymbol{\mu}_{\tilde{v}t}^\top \boldsymbol{\epsilon}_{\tilde{v},t+1}$ captures time-varying risk premia of asset return shocks:[18]

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

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 $\tilde{\boldsymbol{v}}_t$.[19] Hence, defining the *conditional* risk premia analogously as the unconditional ones, we have that the

[18] 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 $m_{t+1} - \mathbb{E}_t[m_{t+1}] = -\boldsymbol{\alpha}^\top \left(\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{\beta}_{\tilde{v}}^\top + \boldsymbol{\Sigma}_{wr}\right)^{-1} \boldsymbol{\beta}_{\tilde{v}}\boldsymbol{\epsilon}_{\tilde{v},t+1} - (\boldsymbol{\lambda}_{\tilde{v}} + \boldsymbol{\mu}_{\tilde{v},t})^\top \boldsymbol{\beta}_{\tilde{v}}^\top \left(\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{\beta}_{\tilde{v}}^\top + \boldsymbol{\Sigma}_{wr}\right)^{-1} \boldsymbol{\beta}_{\tilde{v}}\boldsymbol{\epsilon}_{\tilde{v},t+1}$. Following similar derivations as in Appendix A.2, we can derive that $m_{t+1} - \mathbb{E}_t[m_{t+1}] \to -(\boldsymbol{\lambda}_{\tilde{v}} + \boldsymbol{\mu}_{\tilde{v},t})^\top \boldsymbol{\epsilon}_{\tilde{v},t+1}$ as $N \to \infty$.

[19] This moving average representation is exact only in the continuous time limit. In discrete time, an additional $\boldsymbol{\mu}_{\tilde{v}t}^\top \boldsymbol{\epsilon}_{\tilde{v},t+1}$ term emerges in the MA. We focus on this simplified version for expositional simplicity and because, as shown in Section IA.6 of the Internet Appendix, adding the omitted term the estimated dynamics of $g$ are virtually identical: the omitted term captures less than 1% of the time series variation of macro quantities and the cumulative impulse responses of $g$ to a one-standard-deviation innovations in $\boldsymbol{\mu}_{\tilde{v}t}^\top \boldsymbol{\epsilon}_{\tilde{v},t+1}$ are very small and not significantly different from zero for 18 out of the 21 economic variables. Note also that a *conditional* priced MA representation along the lines of the one in Theorem 1 does hold. This result is available upon request.

time-varying term structure of risk premia is given by

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

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 $\mathbb{E}\left[\boldsymbol{\mu}_{\tilde{v},t-1}\right] = 0$, 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 1.1 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.3). Fourth, to elicit the time variation of the term structure, we need to explicitly model the conditional mean process 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 $q$:

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

The additional layer in equation (22) requires a minimal change to our Gibbs sampler to characterize the posterior distribution. The only deviation from Section 1.1 is that $\boldsymbol{v}_t$ follows a VAR process rather than an IID normal distribution. 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 these parameters follows the usual normal-inverse-Wishart distribution and can be sampled accordingly. We summarize the Gibbs sampler in Proposition A1 of Appendix A.3 and derive it in Internet Appendix IA.1.3.

The time-varying framework in this subsection is closely connected to the literature on affine term structure models. For instance, in Cochrane and Piazzesi (2008) $\boldsymbol{x}_t$ 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 Giglio et al. (2023) in that we study the log SDF linear in latent factors of equity excess returns, impose log normality, and presume that the time-varying risk prices of the shocks to latent factors 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 Giglio et al. (2023) focus on dividend yields. Second, Giglio 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 $g_t$ that always exists, as in equation (20). Our modelling choice is analogous to most macro-finance 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.

# 2 Simulations

We now study the finite-sample properties of our estimator in Proposition 1 via Monte Carlo simulations. We consider two sample sizes, $T \in \{200, 600\}$, matching the quarterly and monthly frequencies, respectively. First, we simulate asset returns from a five-factor model as in equations (1) and (2), as follows $\boldsymbol{r}_t = \hat{\boldsymbol{\alpha}} + \hat{\boldsymbol{\beta}}_{\tilde{v}}\hat{\boldsymbol{\lambda}}_{\tilde{v}} - \frac{1}{2}\hat{\boldsymbol{\Upsilon}}_r + \hat{\boldsymbol{\beta}}_{\tilde{v}}\tilde{\boldsymbol{v}}_t + \boldsymbol{w}_{rt}$, $\tilde{\boldsymbol{v}}_t \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0}_K, \boldsymbol{I}_K)$. Specifically, $\boldsymbol{r}_t$ contain Fama-French 275 portfolio returns (FF275, see Internet Appendix IA.4), 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 $J = \max\{10, \text{int}\big(N/20\big)\}$, $\beta = 0.1$,[20] 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 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:

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

[20] $\beta$ 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.

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 $R_g^2 \in \{30\%, 20\%, 10\%\}$.

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), estimates are biased because $g_t$ 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.

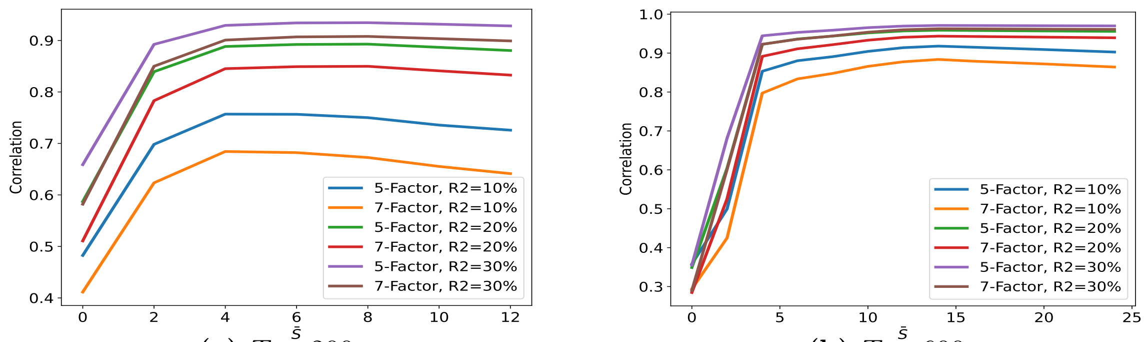
Can we recover the priced information embedded in $g_t$ if we consider only the contemporaneous correlation between $g_t$ and asset returns? To answer this question, we estimate the models with different numbers of lags $\bar{S}$. Figure 1 plots the average correlation between the true $f_t$ and its 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.[21] The model with $\bar{S} = 0$ 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 tends to be a conservative strategy since it delivers proper, yet wider, credible intervals of risk premia estimates. In the empirical application, we explore whether our risk premia estimates are robust to adding more latent factors.

In Tables IA.IV and IA.V of the Internet Appendix, we investigate useless, yet persistent,

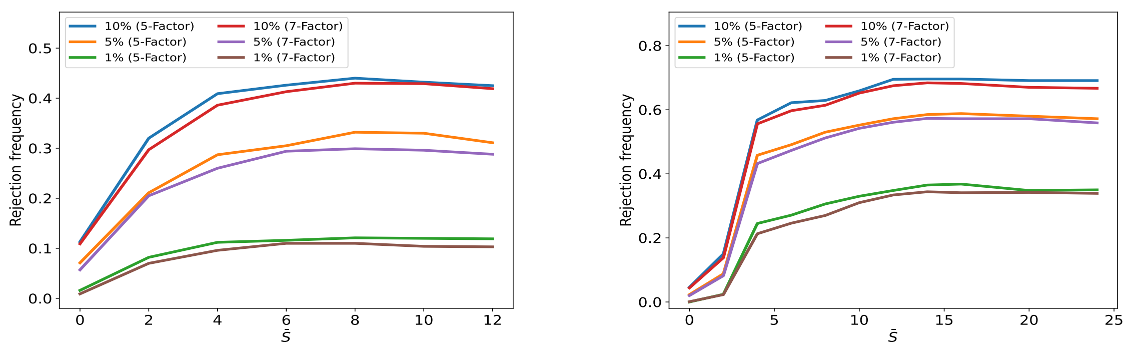
[21] We report the power for $R_g^2 \in \{10\%, 20\%\}$ in Figure IA.1 of the Internet Appendix.




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

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

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




**(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, in 1,000 simulations, of rejecting the null hypothesis $H_0 : \lambda_g^{\bar{S}} = 0$, based on the 90%, 95%, and 99% credible intervals generated by 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 $\bar{S}$.

factors that do not correlate with asset returns. Therein, a larger $R_g^2$ corresponds to a more persistent process. Fama-MacBeth and GMM estimates of useless factor risk premia tend to appear spuriously significant (e.g., Kan and Zhang (1999a,b)). Our Bayesian estimates do not suffer from this issue: The credible intervals of useless factors' risk premia tend to be conservative. Another potential concern is that including many lags of multiple latent factors might lead to severe overfitting of the data. The posterior means of $R_g^2$ in simulations in Table IA.VI of the Internet Appendix show that this is not the case. Furthermore, we explore the performance of our Bayesian estimates for factors that correlate with only the contemporaneous asset return shocks—the setting 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 this special case.

Finally, in Internet Appendix IA.3, we repeat our simulation study to examine the time-varying risk premia and their term structures as described in Section 1.2. Overall, size and power, as well as the correlation between filtered and calibrated latent processes, are similar to those reported in this section. Despite the significant added generality, we observe only a minimal degree of attenuation bias and increased posterior uncertainty.

# 3 Empirical Analysis

In this section, we apply our Bayesian framework to investigate whether factors are priced, the unconditional and time-varying term structure of macroeconomic risk premia, and how these premia connect to the business cycle and forward equity yields.

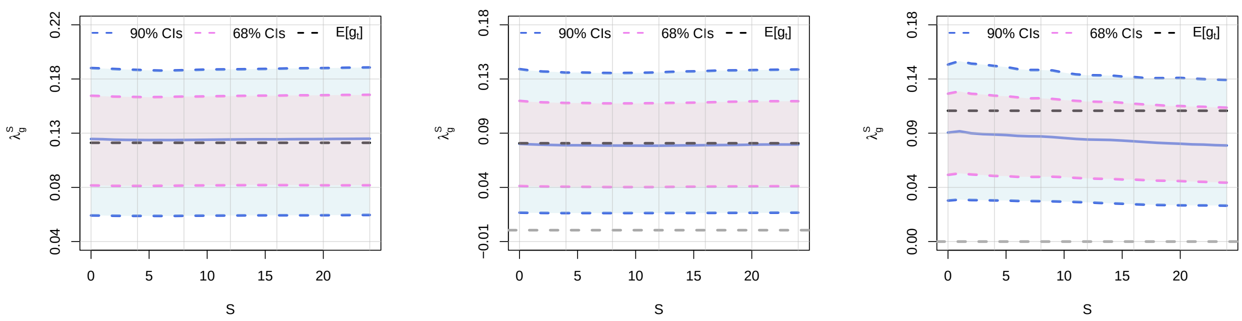
## 3.1 Unconditional Risk Premia

We begin our empirical investigation with unconditional risk premia. Our analysis relies on a large cross-section of FF275, covering the period 1963:Q3–2019:Q4. Throughout our paper, we standardize the tested factors to have unit variances per period. The definition, sample periods, and data sources of factors and test assets can be found in Internet Appendix IA.4.

To conduct our Bayesian estimation of Section 1, we need to determine the number of latent factors, $K$. We adopt the selection approach proposed by Giglio and Xiu (2021) and estimate that the number of factors is five in FF275 at monthly or quarterly frequencies.[22]

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,

[22] 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. Alternatively, one could use the spike-and-slab prior of Bryzgalova et al. (2023) to either select or aggregate factors in the SDF. For robustness, we have re-estimated all the key results of the paper using six- and seven-factor specifications and found virtually identical point estimates.




**(a)** MKT **(b)** SMB **(c)** HML

**Figure 3:** Term structure of risk premia: Fama and French (1993) three factors

Term structure of risk premia (monthly Sharpe ratio units) using Proposition 1, where $\lambda_g^S$ is defined in equation (6). Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns. We estimate risk premia for the Fama and French (1993) three factors using 24-month lags in $g_t$'s equations. Grey dotted lines show in-sample monthly Sharpe ratios. Shaded areas show 68% (pink) and 90% (blue) Bayesian credible intervals. Data sources in Internet Appendix IA.4. Sample: July 1963–December 2019.

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).

### 3.1.1 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 Fama and French (1993) three factors, whose risk premia are estimated using Proposition 1 ($\bar{S} = 24$ 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 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 growth rates in industrial production (IP), GDP, durable and nondurable consumption, dividends, unemployment rate, hours worked, and investment.[23] More interestingly, most of them have *upward-sloping* term

[23]Dividend growth is the quarterly growth of the smoothed aggregate dividend payments made in the previous

**Table 1:** Factors' risk premia

| **Panel A.** Quarterly variables, $\bar{S} = 12$ quarters | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $S =$ | 0 | 2 | 4 | 6 | 8 | 10 | 12 | $R^2_g$ | $R^2_{g,pred}$ | $R^2_{err,pred}$ |
| GDP growth | 0.026* | 0.084*** | 0.133*** | 0.164*** | 0.180*** | 0.195*** | 0.204*** | 25.7% | 23.9% | 9.9% |
| IP growth | 0.008 | 0.087*** | 0.145*** | 0.177*** | 0.194*** | 0.204*** | 0.205*** | 39.0% | 38.8% | 9.7% |
| Durable consumption growth | -0.014 | 0.079** | 0.122*** | 0.140*** | 0.147*** | 0.153*** | 0.158*** | 20.8% | 20.0% | 8.1% |
| Nondurable consumption growth | 0.042*** | 0.103*** | 0.141*** | 0.179*** | 0.206*** | 0.226*** | 0.244*** | 25.1% | 22.3% | 7.1% |
| Unemployment rate change | -0.024 | -0.117*** | -0.203*** | -0.269*** | -0.322*** | -0.366*** | -0.403*** | 47.1% | 46.0% | 9.0% |
| Hours worked growth | 0.024 | 0.095*** | 0.169*** | 0.223*** | 0.261*** | 0.295*** | 0.319*** | 35.9% | 34.7% | 9.0% |
| Investment growth | 0.010 | 0.091*** | 0.158*** | 0.201*** | 0.223*** | 0.240*** | 0.248*** | 36.1% | 35.9% | 5.4% |
| Dividend growth | 0.009 | 0.045* | 0.109*** | 0.175*** | 0.245*** | 0.306*** | 0.357*** | 41.8% | 41.5% | 17.6% |
| Nondurable + service | 0.028* | 0.067* | 0.099* | 0.127** | 0.148* | 0.163** | 0.181** | 19.8% | 17.4% | 23.4% |
| Service consumption growth | 0.006 | 0.013 | 0.020 | 0.027 | 0.035 | 0.041 | 0.045 | 11.0% | 9.7% | 26.5% |
| AEM intermediary | 0.082*** | 0.077** | 0.078** | 0.063 | 0.046 | 0.026 | 0.019 | 16.7% | 11.3% | 5.8% |
| Capital share growth | 0.008 | 0.009 | 0.005 | 0.001 | -0.003 | -0.007 | -0.012 | 9.6% | 8.7% | 15.0% |
| Labor income growth | 0.000 | 0.002 | 0.003 | 0.004 | 0.005 | 0.005 | 0.009 | 7.5% | 7.2% | 8.3% |
| TFP growth | 0.023 | 0.052 | 0.066 | 0.070 | 0.065 | 0.060 | 0.054 | 18.7% | 15.4% | 12.4% |
| TFP growth (util) | 0.000 | 0.000 | 0.000 | -0.001 | -0.002 | -0.001 | 0.000 | 12.2% | 12.0% | 8.1% |
| **Panel B.** Monthly variables, $\bar{S} = 24$ months | | | | | | | | | | |
| $S =$ | 0 | 4 | 8 | 12 | 16 | 20 | 24 | $R^2_{g,total}$ | $R^2_{g,pred}$ | $R^2_{err,pred}$ |
| Nontraded HKM intermediary | 0.098*** | 0.101*** | 0.097*** | 0.093*** | 0.091*** | 0.089*** | 0.088*** | 62.1% | 2.7% | 1.1% |
| Traded HKM intermediary | 0.114*** | 0.115*** | 0.110*** | 0.104*** | 0.100*** | 0.098*** | 0.096*** | 71.8% | 2.7% | 1.1% |
| PS liquidity | 0.050*** | 0.074*** | 0.086*** | 0.097*** | 0.108*** | 0.118*** | 0.126*** | 17.2% | 5.7% | 4.4% |
| $\Delta\log(\text{VIX})$ | -0.131*** | -0.079*** | -0.062*** | -0.049*** | -0.042*** | -0.037*** | -0.032*** | 53.6% | 8.6% | 5.5% |
| Oil price change | -0.004 | -0.023 | -0.034 | -0.039 | -0.042 | -0.041 | -0.040 | 9.5% | 9.0% | 10.3% |
| TED spread change | 0.000 | -0.001 | 0.000 | 0.000 | 0.001 | 0.001 | 0.001 | 11.0% | 10.2% | 15.0% |

Bayesian estimates of factors' risk premia using Proposition 1, where $\lambda_g^S$ is defined in equation (6). Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns. Panel A: quarterly factor risk premia using 12-quarter lags in $g_t$'s equations. Panel B: monthly factor risk premia using 24-month lags. Significance based on Bayesian credible intervals: * (**, ***) indicates the 90% (95%, 99%) interval excludes zero. Data sources in Internet Appendix IA.4. $R^2_g$: total $R^2$ from priced shocks; $R^2_{g,pred}$: predictive $R^2$ from priced shocks (excluding contemporaneous shock); $R^2_{err,pred}$: predictive $R^2$ from error term $w_{gt}$ via AR(12).

structures of risk premia. At the 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, macro factors are riskier from the perspective of long-term than short-term investors. Interestingly, despite this common pattern of sharply increasing term structures of risk premia for many macroeconomic factors, we find that total factor productivity (TFP) is not significantly priced at any horizon. As discussed in Section 3.1.2 below, the commonality is driven by these macro factors loading on essentially the exact same priced shock—but not the TFP one. This presents a novel challenge for equilibrium macro finance models.

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

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.

and Cochrane (1999)) and long-run risk frameworks (Bansal and Yaron (2004)).[24] 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 for smoothed dividend growth, the magnitudes and slopes are much more sizable in the long-run risk model. Overall, the long-run risk model tends to be more in line with our estimates for nondurable consumption and dividend growth.

Third, priced shocks explain a large share of the time series variance (see $R^2_g$) of the priced macro variables and capture most of their predictability (see $R^2_{g,pred}$ vs $R^2_{err,pred}$). This degree of commonality between macro and financial quantities is significantly larger and more sharply identified than what is normally measured using traditional approaches (e.g., the link between consumption and returns is weakly identified at best in Kleibergen and Zhan (2020)). Our identification strategy, contrary to the standard approaches, recovers the priced innovations from a large cross-section of asset returns rather than simply the standalone macro variables or their AR(1) residuals. The last three columns of Table 1 reveal why the standard approach does not properly capture the connection between macro and finance: priced and unpriced components of macro time series have quite different degrees of persistency. For example, priced shocks explain more than a quarter of the time series variation of GDP growth, with a predictive $R^2$ of about 24%, while unpriced shocks are also predictable albeit to a much lower extent (with a predictive $R^2$ for this component of about 10%). Not distinguishing these two components and just focusing on AR(1) residuals in macro variables leads to inconsistent estimates that suffer from attenuation bias.

Fourth, 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 shown 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.

Moreover, the term structure of VIX risk premia (more precisely, their absolute values)

[24] We discuss the calibrations in detail in Internet Appendix IA.5.

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 verify our baseline risk premia estimates by directly projecting $g$ onto the space of the SDF shocks: $g_t = \mu_g + \sum_{s=0}^{\bar{S}} \rho_s m_{t-s} + w_{gt}$. The term structure of risk premia is then:

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

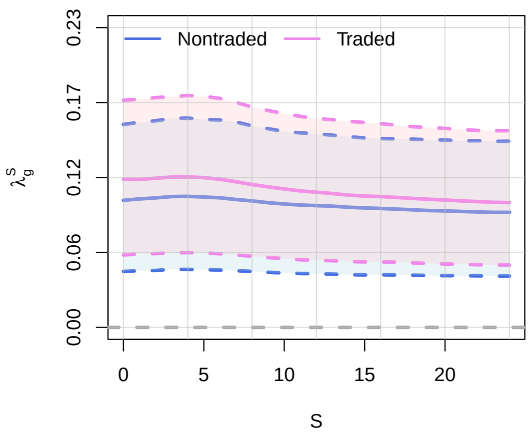
Figure IA.3 of the Internet Appendix presents the term structure of risk premia estimates under this restricted model. We show that most macro factors that have been identified as being priced in the unrestricted model display similar patterns in both the magnitudes of risk premia estimates and shapes of their term structures. In other words, not imposing the restriction ($\boldsymbol{\eta_g} \propto \boldsymbol{\lambda_v}$) barely affects the risk premia estimates. However, as the time series fit is reduced, it leads to wider estimation uncertainty and wider credible intervals. It is also worth noting that TFP growth, especially the utilization-adjusted measure, stays always unpriced.[25]

However, canonical mimicking portfolios based on single-period risk exposures may fail to capture the entire term structure of risk premia embedded in economic factors. Figure 4 plots 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 almost zero "forward betas" (see 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.

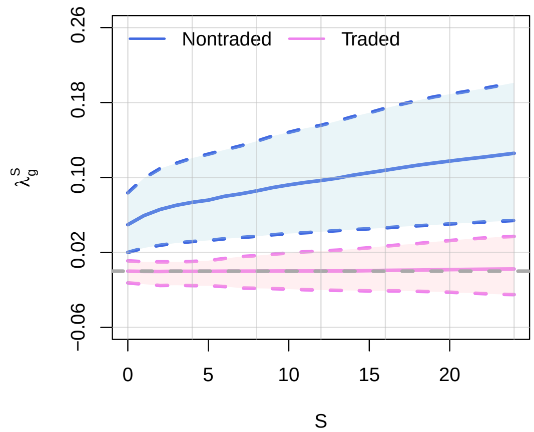
### 3.1.2 Commonality and conditional means of Macro Factors

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? Is the commonality among macro variables caused by their exposure to the same priced shocks? This is what Theorem 1 implies, and we

[25] We further confirm that we can interpret the term structure of risk premia estimates from the angle of horizon-specific mimicking portfolios. As we find in unreported results available upon request, these portfolios display increasing term structures of risk premia that are similar to what we find in Table 1.




**(a)** HKM intermediary factors




**(b)** PS liquidity factors

**Figure 4:** Term structure of factor's risk premia: Traded vs. Nontraded versions

Term structure of monthly risk premia for traded and nontraded versions of He et al. (2017) intermediary factors and Pástor and Stambaugh (2003) liquidity factors. Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns. Point estimates shown with shaded areas indicating 90% Bayesian credible intervals. Data sources in Internet Appendix IA.4.

**Table 2:** Are the MA components of macro factors similar?

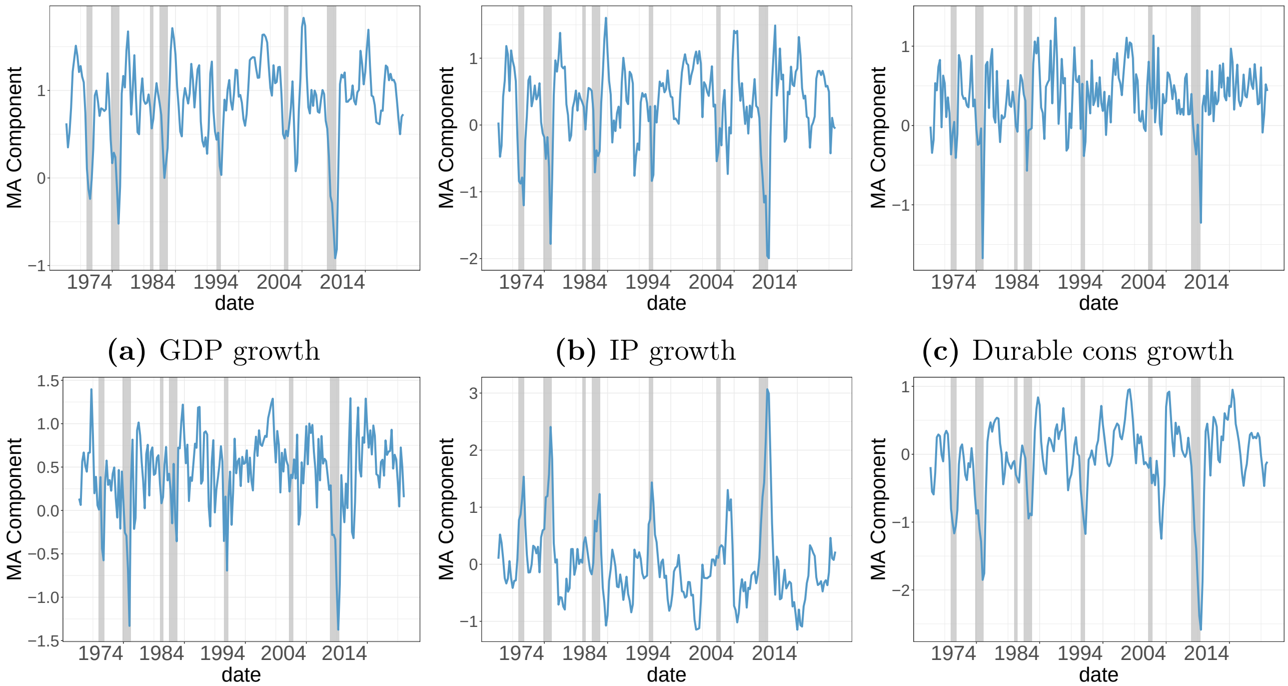
| | GDP growth | IP growth | Durable | Nondurable | Service | Dividend growth | Unemployment | Hours worked | Investment |
|---|---|---|---|---|---|---|---|---|---|
| GDP growth | 1.00 | 0.90 | 0.69 | 0.70 | 0.60 | 0.39 | -0.86 | 0.90 | 0.95 |
| IP growth | 0.90 | 1.00 | 0.72 | 0.70 | 0.57 | 0.35 | -0.83 | 0.82 | 0.90 |
| Durable | 0.69 | 0.72 | 1.00 | 0.64 | 0.35 | 0.32 | -0.66 | 0.63 | 0.70 |
| Nondurable | 0.70 | 0.70 | 0.64 | 1.00 | 0.59 | 0.48 | -0.76 | 0.64 | 0.60 |
| Service | 0.60 | 0.57 | 0.35 | 0.59 | 1.00 | 0.43 | -0.62 | 0.60 | 0.52 |
| Dividend growth | 0.39 | 0.35 | 0.32 | 0.48 | 0.43 | 1.00 | -0.72 | 0.65 | 0.39 |
| Unemployment | -0.86 | -0.83 | -0.66 | -0.76 | -0.62 | -0.72 | 1.00 | -0.97 | -0.86 |
| Hours worked | 0.90 | 0.82 | 0.63 | 0.64 | 0.60 | 0.65 | -0.97 | 1.00 | 0.90 |
| Investment | 0.95 | 0.90 | 0.70 | 0.60 | 0.52 | 0.39 | -0.86 | 0.90 | 1.00 |

Correlation among moving average components spanned by latent factors, $\sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top \boldsymbol{v}_{t-s}$, with $\bar{S} = 12$ quarters. Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns. Data sources in Internet Appendix IA.4.

now turn to verifying this hypothesis.

Figure 5 plots the posterior means of the MA components, $\sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top \boldsymbol{v}_{t-s}$, of six priced macro variables. Strikingly, the *priced* MA components of all these variables present clear business cycle patterns—long-horizon investors who hedge against low (high) realizations of these macro factors require positive (negative) risk premia. As shown in Table 2, GDP, IP, unemployment, hours worked, and investment have highly correlated MA components, often with coefficients of about 90%. Instead, the MA components of variables that tends *not* to display significant risk premia, although correlated, seem to contain independent information.

But what drives the commonality? As we are about to show, several macro variables load




**(a)** GDP growth **(b)** IP growth **(c)** Durable cons growth

**(d)** Nondurable cons growth **(e)** Unemployment rate change **(f)** Hours worked growth

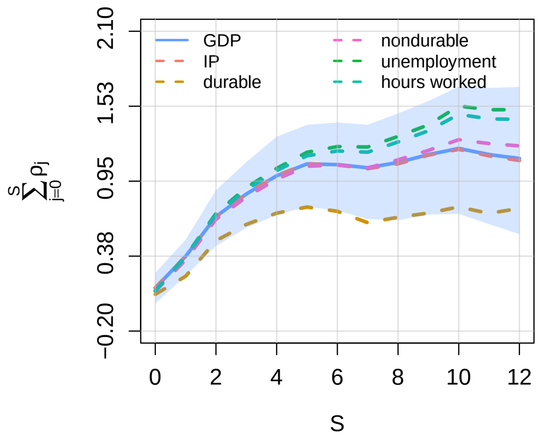
**Figure 5:** Moving average components of some macro factors

Time series of moving average components (posterior means) spanned by latent factors: $\sum_{s=0}^{\bar{S}} \rho_s \boldsymbol{\eta}_g^\top \boldsymbol{v}_{t-s}$, with $\bar{S} = 12$ quarters. Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns. Data sources in Internet Appendix IA.4. Sample: 1963Q3–2019Q4.

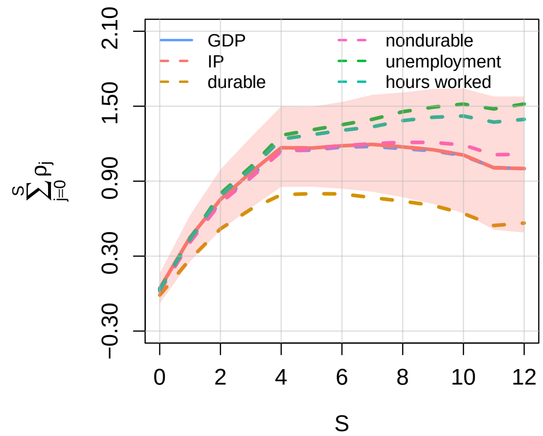
on almost identical priced shocks—as our Theorem 1 implies. To illustrate this point, consider the following exercise. Suppose that $\hat{f}_t$ is the financial shock identified as the driver of the (priced) conditional mean of $g_t$.[26] We can then elicit the responses of *other* macro variables (e.g., GDP, consumption, unemployment, etc.) to the same shock. If the financial shocks underlying different macro variables are interchangeable, we should observe roughly the same impulse responses to these financial shocks for all of them.

In Figure 6, we report six macro variables: GDP, IP, durable and nondurable consumption, unemployment, and hours worked. For each variable, we estimate its impulse responses to six different $f_t$ shocks identified by targeting the individual factors. Shaded areas denote 90% confidence bands of CIRFs of each variable to its own $f_t$ shock. The figure shows that the financial shocks are almost interchangeable, except possibly for the one of durable consumption.

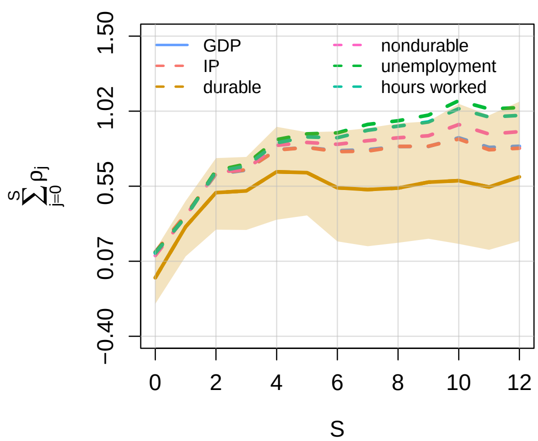
[26] Note that our identification approach finds the best linear combination of financial shocks $\boldsymbol{v}_t$, $f_t = \boldsymbol{\eta}_g^\top \boldsymbol{v}_t$, such that $\{f_{t-s}\}_{s=0}^{\bar{S}}$ explain the maximal fraction of time-series variation in $g_t$. That is, our method echoes the max-share identification strategy of Faust (1998), Uhlig (2003), Barsky and Sims (2011), Francis et al. (2014), Angeletos et al. (2020).



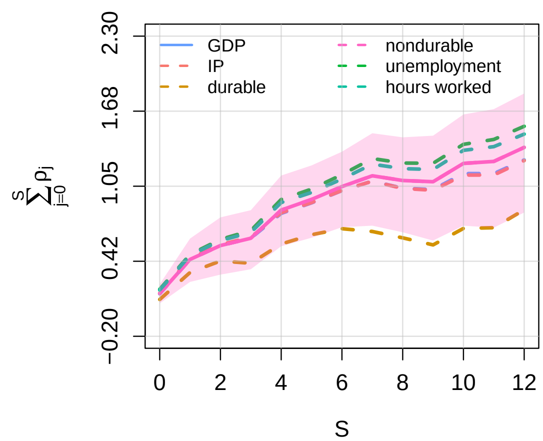

**(a)** GDP growth



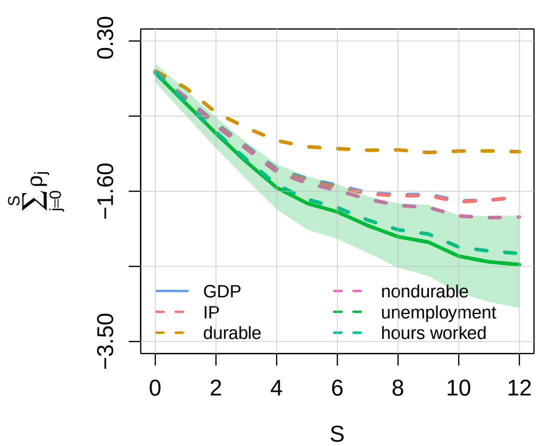

**(b)** IP growth



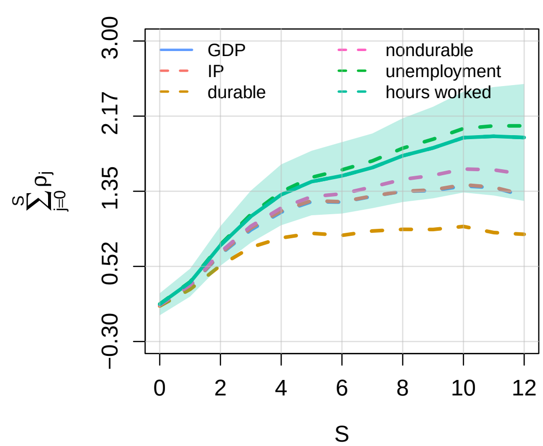

**(c)** Durable cons. growth




**(d)** Nondurable cons. growth




**(e)** $\Delta$unemployment rate$_t$




**(f)** $\Delta\log(\text{hours worked}_t)$

**Figure 6:** Cumulative IRFs to Different Financial Shocks

Impulse-response functions of macro time series to priced shocks driving their priced conditional means. Panel (a): responses to the GDP priced shock; panel (b): responses to the IP priced shock, etc. Shaded areas show posterior 90% confidence regions for each variable's response to its own shock. Base assets: 275 Fama-French characteristic-sorted portfolios. Data sources in Internet Appendix IA.4.

The evidence suggests that an (almost) identical financial shock drives the dynamics of multiple macro variables at horizons of two to three years, and the commonality in the priced conditional means comes from a common response pattern among macro variables to priced financial shocks. In most models, such commonality could arise due to the all-encompassing effects of total factor productivity (Angeletos et al. (2020)). Nevertheless, as shown in Table 1, TFP shocks are not priced at any horizon, presenting a new challenge for equilibrium macro-finance models.

### 3.1.3 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. That is, we quantify the saliency of priced macro shocks in financial markets. We use the SDF representation in Remark 1; that is, $m_t = \kappa_m - \lambda_f f_t - \boldsymbol{\lambda}_u^\top \boldsymbol{\Sigma}_u^{-1} \boldsymbol{u}_t$, where $\boldsymbol{u}_t$ are orthogonal to $f_t$ and act as the control for omitted sources of priced risk. Table 3 reports the risk price estimates of $f_t$ for several priced nontraded risk factors based on the evidence in Table 1. These $f_t$ shocks are indeed significantly priced in the cross-section with extremely similar risk prices (in Panel A)—as our priced Wold representation theorem implies—once again suggesting a common origin for the risk compensation demanded by these macro quantities. Furthermore, the annualized 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 macroeconomic sources of risk explain individually about $13 - 58\%$ of the SDF's variance.

Our findings indicate that macroeconomic risk is of first-order importance in financial markets. Yet, as the last column of Table 3 highlights, a significant amount of priced shocks (half, or more, of the total Sharpe ratio achievable in the economy) are *not* captured by these economic factors. This once again emphasizes that using the standalone macro variables, or their AR(1) innovations, to connect macro to finance via canonical asset pricing exercises (e.g., Fama-MacBeth regressions), is problematic without explicitly controlling for omitted variables (Giglio and Xiu (2021)).

### 3.1.4 Contemporaneous Innovations in Macro Factors

Empirically, researchers often fail to identify priced macro risk when studying only contemporaneous correlations of returns and factors. The first column of Table 1 indicates that the risk premia of many macro factors are minimal and often statistically insignificant at $S = 0$. One concern about these results is that we include many lags in our specification, leading to poten-

**Table 3:** Risk price of the $f_t$ shock to nontraded risk factors

| | $\lambda_f$ | $\mathbb{E}[SR_f \mid \text{data}]$ | $\mathbb{E}\left[\frac{SR_f^2}{SR_m^2} \mid \text{data}\right]$ |
|---|---|---|---|
| **Panel A.** Quarterly variables, $\bar{S} = 12$ quarters | | | |
| 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] |
| Unemployment rate change | -0.241 [-0.361, -0.123] | 0.483 [0.246, 0.722] | 0.262 [0.076, 0.511] |
| Hours worked growth | 0.243 [0.119, 0.363] | 0.487 [0.238, 0.726] | 0.264 [0.073, 0.527] |
| Investment growth | 0.253 [0.126, 0.371] | 0.507 [0.253, 0.742] | 0.280 [0.083, 0.546] |
| AEM intermediary | 0.356 [0.205, 0.494] | 0.711 [0.412, 0.987] | 0.580 [0.225, 0.858] |
| **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] |

Risk price of $f_t$ shock to nontraded factors (column $\lambda_f$), annualized Sharpe ratio from the $\lambda_f f_t$ component (column $\mathbb{E}[SR_f \mid \text{data}]$), and share of SDF variance explained by $f_t$ (column $\mathbb{E}\left[SR_f^2/SR_m^2 \mid \text{data}\right]$). Estimates from Table 1 and SDF representation in Remark 1. Each column shows posterior median and 90% credible intervals.

tially noisier risk premia estimates. To alleviate this concern, we repeat the estimation using $\bar{S} = 0$. Panel A of Table IA.X of the Internet Appendix (IA) shows that these macro factors, which are related to quarterly production, consumption, investment, and labour market conditions, carry tiny and largely insignificant risk premia. This is *not* a byproduct of our method, as we uncover the same finding using the frequentist estimator of Giglio and Xiu (2021). Panel B further extracts the AR(1) innovations in macro factors and estimates their risk premia by setting $\bar{S} = 0$. Consistent with the results in Panel A, most macro factors exhibit negligible and insignificant risk premia (a finding confirmed using frequentist inference). Although the AR(1) model is commonly used in both empirical and theoretical works, extracting 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. Instead, our MA representation does not take a stance on the exact data-generating processes, and yields a flexible and robust function of both 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 IA.X. For instance, the contemporaneous asset return shocks explain only 2% of time series variations in investment growth, but its $R_g^2$ increases to 36% in the estimation with $\bar{S} = 12$ quarters, hence greatly enhancing the signal-to-noise ratio and our ability to identify the risk premia.

The contrast between the risk premia of most macro factors in Tables 1 and IA.X is striking: seemingly unpriced macro risk at high frequency becomes first-order at business cycle frequency. There is extensive literature on developing new estimators of risk premia that are robust to weak factors—factors that have small, or vanishing, correlations with asset returns, yielding a first-order identification problem for their risk premia.[27] Our estimator is not only robust to the weak identification issue but, more importantly, uncovers that many macro factors that appear weak at short horizons are indeed strongly identified once their persistence is properly accounted for using our MA-based approach. This allows us to correctly identify the priced innovations of macro factors, trace their propagation mechanism, and in turn uncover the very large risk premia that they command at business cycle frequency.

The previous literature has produced conflicting empirical evidence on priced consumption risk at different frequencies: from insignificant and weakly identified at quarterly horizons, to strongly priced at business cycle horizons.[28] Our results not only indicate that there is no contradiction among these different estimates, but also uncover that all of them are generated by the multi-period responses of macro variables to priced financial shocks. That is, our approach not only identifies the entire term structure of risk premia, but also its propagation mechanism.

## 3.2 Time-Varying Term Structure of Macroeconomic Risk Premia

We now turn to the analysis of the time variation in the term structure of macro factors' risk premia, applying the method in Section 1.2. 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-earnings ratio, as well as term, default, and value spreads. Using this formulation, we estimate the term structure of unconditional risk premia for the same set of variables as in Table 1.

Figure IA.4 of the Internet Appendix shows the empirical results. The blue lines and shaded areas present the estimates based on the conditional models. For comparison, we also

[27]See, e.g., Kan et al. (2013), Gospodinov et al. (2014, 2019), Kleibergen and Zhan (2020), Anatolyev and Mikusheva (2022), and Bryzgalova et al. (2023).

[28]See, e.g., the disagreement among estimates for consumption risk premia Lettau and Ludvigson (2001), Jagannathan and Wang (2007), Hansen et al. (2008), Ortu et al. (2013), Kleibergen and Zhan (2020), and Bandi and Tamoni (2023).

include the previous estimates (the 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. Overall, the estimates of risk premia based on the unconditional specification appear consistent 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.[29]

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

## 3.3 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 the past literature (e.g., van Binsbergen et al. (2012), Bansal et al. (2021), and Giglio et al. (2023)).

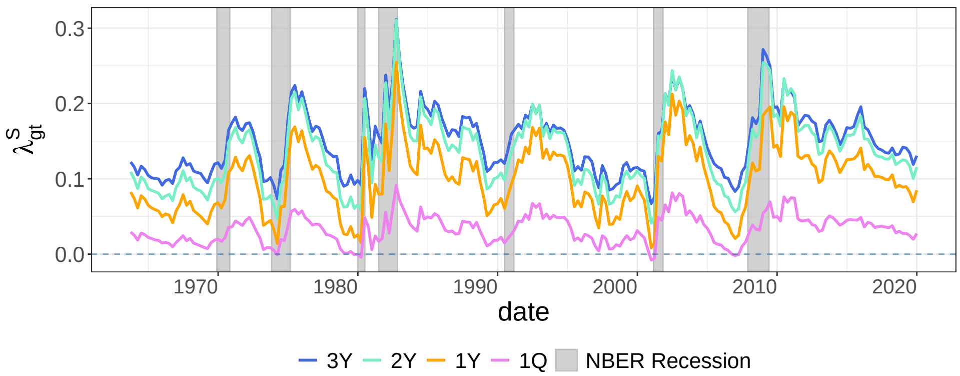
In Internet Appendix IA.7, we show that under joint log-normality of SDF and dividend growth, forward equity yield ($e^f$) and dividend risk premia satisfy the following relationship:

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

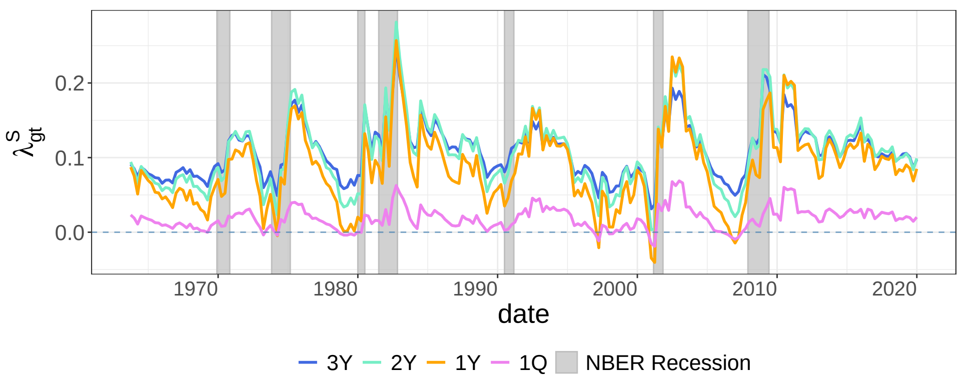
where $g_{d,t,t+s} = \frac{1}{s}\log\left(\frac{D_{t+s}}{D_t}\right)$ is the per-period log dividend growth rate, $\Delta d_{t,t+s} = \log\left(\frac{D_{t+s}}{D_t}\right)$ is the multiperiod dividend growth, and $\lambda^s_{dt} = -\frac{1}{s}\text{cov}_t(m_{t,t+s}, \Delta d_{t,t+s})$ is the $s$-period 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.7, dividend risk premia can be interpreted as the per-period risk premium on the hold-to-maturity dividend strips. We estimate

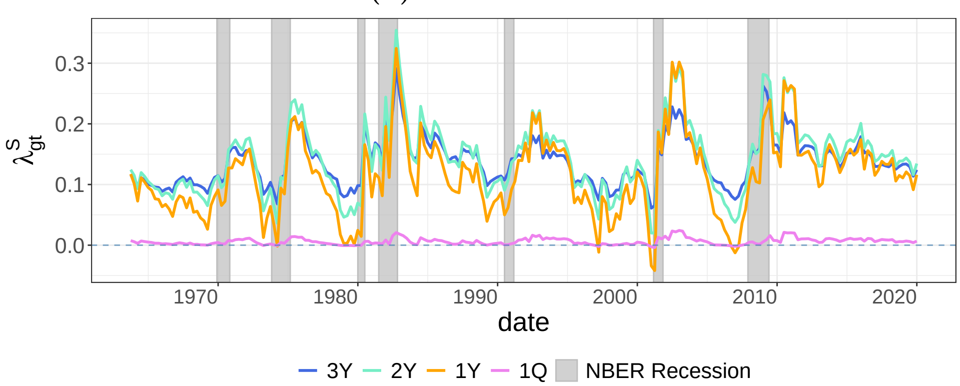
[29] We obtain similar results for durable consumption, dividends, hours worked, and unemployment rate.




**(a)** Nondurable Consumption Growth




**(b)** GDP Growth




**(c)** Industrial Production Growth

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

Time-varying term structure of quarterly risk premia following Section 1.2. Risk premia of latent factors are linear in four external predictors: PE ratio of S&P 500, term spread, default spread, and value spread. Base assets: 275 Fama-French characteristic-sorted portfolios. Data sources in Internet Appendix IA.4.

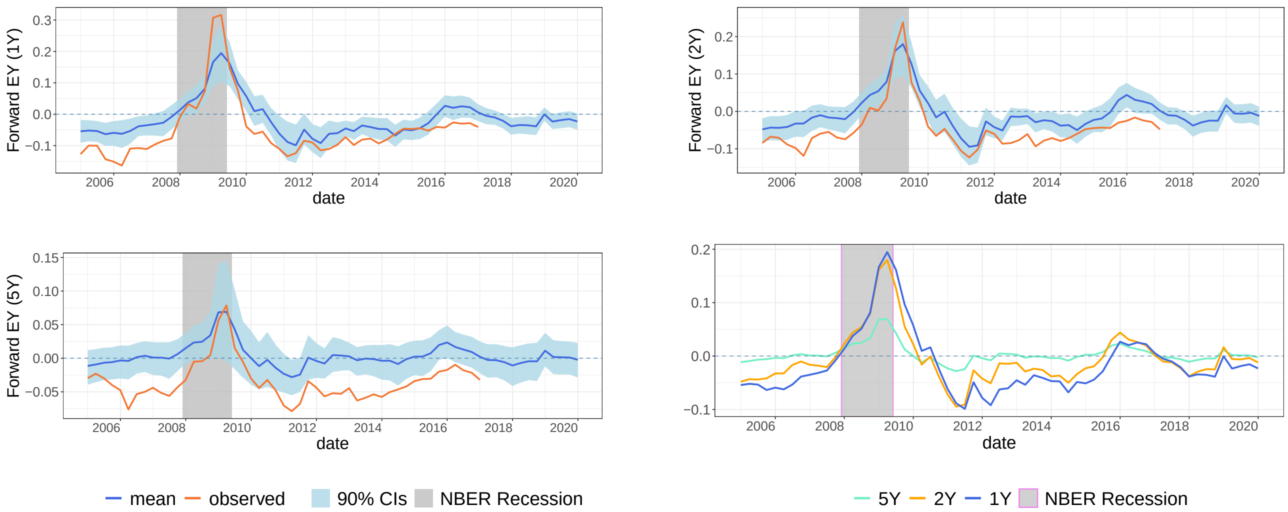



**Figure 8:** Time series of estimated forward equity yields

Time series of estimated forward equity yields based on our MA model with 20 lags. Dividend risk premia are time-varying and linearly dependent on external predictors. We estimate a five-factor latent factor model using FF275 for 1963Q3–2019Q4 but plot only 2004Q4–2019Q4, the period with observed forward equity yield data. The bottom right panel shows posterior means for all horizons considered. Observed forward equity yields are from Bansal et al. (2021).

the term structure of unconditional dividend risk premia using our MA formulation and show in Figure IA.2 of the Internet Appendix 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.

To obtain the term structure of dividend strips, we need to estimate the conditional mean of dividend growth. We proxy the conditional mean using equation (5) 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 (5) the unspanned component $w_{gt}$ is allowed to be persistent. In other words, we do not assume that the MA component of priced shocks captures the *entire* dividend predictability. Consequently, the forward equity yields implied by our priced MA representation are not guaranteed to match the empirically observed ones exactly.

We estimate forward equity yields for one-, two-, and five-year holding horizons using equation (25). The estimation uses the full sample (1963Q3–2019Q4), but Figure 8 displays only 2004Q4–2019Q4, the period with observed forward equity yield data.[30] Figure 8 shows that our

[30]The data on realized one-, two-, and five-year forward equity yields are from Bansal et al. (2021). We thank the authors for sharing the data with us.

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. The bottom right panel reports our estimates; remaining panels show observed yields alongside our estimates and 90% credible intervals. Our estimates are also strongly countercyclical, closely tracking the observed data. As we point out in equation (26), our MA formulation allows for other sources of *unpriced* dividend predictability that we leave unmodelled, likely driving the relatively minor differences between our estimates and the observed data.

# 4 Conclusion

Based on a novel representation result, we propose a new 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 while, following our priced Wold decomposition, the tested factor can slowly adjust to the asset return systematic shocks. This allows the tested factors and asset returns to have rich dynamics but poses a challenge for frequentist estimation. We tackle this challenge by taking a Bayesian perspective. Specifically, we derive a Gibbs sampler in which all conditional distributions of model parameters have standard closed forms, making our estimator straightforward to implement.

We apply our method to a large cross-section of equity portfolio returns. Unconditionally, most macro variables have significantly upward-sloping term structures of risk premia. Although they seem almost unpriced at quarterly horizons, their risk premia 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 return predictors used in previous literature, macro risk premia are strongly time-varying and have a clear business cycle pattern: They are countercyclical, with low risk premia in normal times but sharply increasing premia in recessions.

Theoretical models predict which economic states and variables should be priced. Our new empirical facts on the term structure and time variation of macro risk premia present a new challenge for equilibrium macro-finance models.

# References


Abel, A. B., J. C. Eberly, and S. Panageas (2013). Optimal inattention to the stock market with information costs and transactions costs. *Econometrica 81*(4), 1455–1481.

Adrian, T., N. Boyarchenko, and D. Giannone (2019). Vulnerable growth. *American Economic Review 109*(4), 1263–1289.

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

Aït-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics 74*(3), 487–528.

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.

Andries, M., T. M. Eisenbach, and M. C. Schmalz (2024, November). Horizon-dependent risk aversion and the timing and pricing of uncertainty. *The Review of Financial Studies 37*(11), 3272–3334.

Angeletos, G.-M., F. Collard, and H. Dellas (2020). Business-cycle anatomy. *American Economic Review 110*(10), 3030–3070.

Backus, D., N. Boyarchenko, and M. Chernov (2018). Term structures of asset prices and returns. *Journal of Financial Economics 129*(1), 1–23.

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

Bańbura, M., D. Giannone, and L. Reichlin (2010). Large Bayesian vector auto regressions. *Journal of Applied Econometrics 25*(1), 71–92.

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., D. Kiku, and A. Yaron (2016). Risks for the long run: Estimation with time aggregation. *Journal of Monetary Economics 82*, 52–69.

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.

Barsky, R. B. and E. R. Sims (2011). News shocks and business cycles. *Journal of Monetary Economics 58*(3), 273–289.

Belo, F. and X. Li (2023). Production-based stochastic discount factors. *Available at SSRN 4287502*.

Bernanke, B. S. and K. N. Kuttner (2005). What explains the stock market's reaction to federal reserve policy? *The Journal of Finance 60*(3), 1221–1257.

Bianchi, F., S. C. Ludvigson, and S. Ma (2022, May). A structural approach to high-frequency event studies: The fed and markets as case history. Working Paper 30072, National Bureau of Economic Research.

Black, F., M. C. Jensen, and M. Scholes (1972). The capital asset pricing model: Some empirical tests. In M. C. Jensen (Ed.), *Studies in the Theory of Capital Markets*, pp. 79–121. New York: Praeger Publishers.

Breeden, D. T., M. R. Gibbons, and R. H. Litzenberger (1989). Empirical test of the consumption-oriented CAPM. *Journal of Finance 44*(2), 231–262.

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 (2025). Consumption in asset returns. *Journal of Finance*. Forthcoming.

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

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.

Chen, N.-F., R. Roll, and S. A. Ross (1986). Economic forces and the stock market. *Journal of Business 59*(3), 383–403.

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.

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*.

Cogley, T. and T. J. Sargent (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics 8*(2), 262–302.

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

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016, March). Parameter learning in general equilibrium: The asset pricing implications. *American Economic Review 106*(3), 664–98.

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.

Croce, M. M. (2014). Long-run productivity risk: A new hope for production-based asset pricing? *Journal of Monetary Economics 66*, 13–31.

Crump, R. K., S. Eusepi, D. Giannone, E. Qian, and A. M. Sbordone (2021, August). A large bayesian var of the united states economy. Staff Report 976, Federal Reserve Bank of New York.

Del Negro, M. and G. E. Primiceri (2015). Time-varying structural vector autoregressions and monetary policy: A corrigendum. *Review of Economic Studies 82*(4), 1342–1345.

Del Negro, M. and F. Schorfheide (2004). Priors from general equilibrium models for VARs. *International Economic Review 45*(2), 643–673.

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*.

Fama, E. F. (1990). Stock returns, expected returns, and real activity. *The Journal of Finance 45*(4), 1089–1108.

Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *The Journal of Finance 47*(2), 427–465.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Faust, J. (1998). The robustness of identified var conclusions about money. In *Carnegie-Rochester conference series on public policy*, Volume 49, pp. 207–244. Elsevier.

Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. *Working paper*.

Francis, N., M. T. Owyang, J. E. Roush, and R. DiCecio (2014). A flexible finite-horizon alternative to long-run restrictions with an application to technology shocks. *Review of Economics and Statistics 96*(4), 638–647.

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

Giannone, D., M. Lenza, and G. E. Primiceri (2015). Prior selection for vector autoregressions. *Review of*

*Economics and Statistics 97*(2), 436–451.

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.

Gomez Cram, R., H. Kung, and H. Lustig (2025, March). Can treasury markets add and subtract? NBER Working Paper 33604, National Bureau of Economic Research. Available at SSRN: https://ssrn.com/abstract=5190791.

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.

Hansen, L. P. and S. F. Richard (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica 55*(3), 587–613.

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.

Herskovic, B., A. Moreira, and T. Muir (2019, January). Hedging risk factors. Technical report, University of California, Los Angeles (UCLA) - Anderson School of Management; University of Rochester - Simon Business School. SSRN Working Paper No. 3148693.

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.

Jordà, Ò., M. Schularick, and A. M. Taylor (2017). Macrofinancial history and the new business cycle facts. In *NBER Macroeconomics Annual 2016, Volume 31*, pp. 213–263. University of Chicago Press.

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.

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

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the fed funds futures market. *Journal of Monetary Economics 47*(3), 523–544.

Kydland, F. E. and E. C. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica 50*(6), 1345–1370.

Lamont, O. A. (2001). Economic tracking portfolios. *Journal of Econometrics 105*(1), 161–184.

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.

Mehra, R. and E. C. Prescott (1985, Mar). The equity premium: A puzzle. *Journal of Monetary Economics 15*(2), 145–161.

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.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies 72*(3), 821–852.

Shanken, J. (1992). On the estimation of beta-pricing models. *Review of Financial Studies 5*(1), 1–33.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review 36*(5), 975–1000.

Sims, C. A. and T. Zha (2006, March). Were there regime switches in u.s. monetary policy? *American Economic Review 96*(1), 54–81.

Smets, F. and R. Wouters (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review 97*(3), 586–606.

Stock, J. H. and M. W. Watson (2011). Dynamic factor models. In M. P. Clements and D. F. Hendry (Eds.), *Oxford Handbook of Economic Forecasting*, pp. 35–59. Oxford, UK: Oxford University Press.

Uhlig, H. (2003). What moves gnp? *Unpublished*.

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

Wilcox, D. W. (1992). The construction of us consumption data: Some facts and their implications for empirical work. *American Economic Review*, 922–941.

Zviadadze, I. (2021). Term structure of risk in expected returns. *Review of Financial Studies 34*(12), 6032–6086.

# Appendices

## A.1 Proof of Theorem 1

We prove the theorem below and report an intuitive construction in Internet Appendix IA.1.1.

*Proof of Theorem 1.* By the Wold theorem for $\{\tilde{m}_t\}$, $\mathcal{M}_t = \mathcal{H}_t \oplus \mathcal{V}$, $\mathcal{H}_t := \overline{\text{span}}\{\varepsilon_s^m : s \leq t\}$, $\mathcal{V} := \bigcap_{n\geq 1} \mathcal{M}_{t-n}$, where $\mathcal{V}$ denotes the deterministic subspace (perfectly predictable from the remote past). Decompose the mean-zero part of $g_t$ as $g_t - \mu_g = \tilde{g}_t + \tilde{w}_t^g$, with $\tilde{g}_t \in \mathcal{M}_t$ and $\tilde{w}_t^g \perp \mathcal{M}_t$. Apply the Wold decomposition to the priced component $\tilde{g}_t$: $\tilde{g}_t = \tilde{g}_t^{ND} + \tilde{g}_t^D$, $\tilde{g}_t^{ND} \in \mathcal{H}_t$, $\tilde{g}_t^D \in \mathcal{V}$. Since $\{\varepsilon_{t-j}^m\}_{j\geq 0}$ is an orthogonal basis of $\mathcal{H}_t$, expand

$$\tilde{g}_t^{ND} = \sum_{j=0}^{\infty} \theta_j \, \varepsilon_{t-j}^m, \qquad \theta_j = \frac{\mathbb{E}[\tilde{g}_t^{ND} \, \varepsilon_{t-j}^m]}{\sigma_\varepsilon^2}.$$

Because both $\tilde{g}_t^D$ and $\tilde{w}_t^g$ are orthogonal to all $\varepsilon_{t-j}^m$, we also have $\theta_j = \frac{\mathbb{E}[(g_t-\mu_g)\,\varepsilon_{t-j}^m]}{\sigma_\varepsilon^2}$.

Parseval's identity implies $\mathbb{E}[(\tilde{g}_t^{ND})^2] = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \theta_j^2 < \infty$, so $\{\theta_j\} \in \ell^2$. Setting $w_t^g := \tilde{g}_t^D + \tilde{w}_t^g$ (collecting the deterministic and unpriced components) yields the expression in equation (4). Uniqueness follows from the uniqueness of the innovation process $\varepsilon_t^m = \tilde{m}_t - \Pi_{\mathcal{M}_{t-1}}(\tilde{m}_t)$ and the orthogonality of the innovation basis. $\square$

## A.2 Lemma A1 and its proof

**Lemma 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, $\alpha_i$ is IID and cross-sectionally independent of factor loadings, with a zero mean and satisfying that $\frac{\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}_{wr}^{-1} \boldsymbol{\beta}_{\tilde{v}}}{N} \to \mathbf{0}_K^\top$ as $N \to \infty$. All elements in $\boldsymbol{\beta}_{\tilde{v}}$ are bounded.*[31]

*Proof.* Assumptions in equation (2) imply that

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

[31] 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 Lemma 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).

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

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

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

Assumption (ii) 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), $\frac{\boldsymbol{\alpha}^\top\boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}}{N} \to \boldsymbol{0}_K^\top$ as $N \to \infty$, which implies that $(I) \to \boldsymbol{0}_K^\top$.

$$(II) = \boldsymbol{\lambda}_{\tilde{v}}^\top\boldsymbol{\beta}_{\tilde{v}}^\top\left[\boldsymbol{\Sigma}_{wr}^{-1} - \boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}(\boldsymbol{I}_K + \boldsymbol{\beta}_{\tilde{v}}^\top\boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}})^{-1}\boldsymbol{\beta}_{\tilde{v}}^\top\boldsymbol{\Sigma}_{wr}^{-1}\right]\boldsymbol{\beta}_{\tilde{v}} = \boldsymbol{\lambda}_{\tilde{v}}^\top\left[\boldsymbol{I}_K - (\boldsymbol{A} + \boldsymbol{I}_K)^{-1}\right]$$

where $\boldsymbol{A} = \boldsymbol{\beta}_{\tilde{v}}^\top\boldsymbol{\Sigma}_{wr}^{-1}\boldsymbol{\beta}_{\tilde{v}}$. Since the eigenvalues of $\boldsymbol{\beta}_{\tilde{v}}^\top\boldsymbol{\beta}_{\tilde{v}}$ will explode as $N \to \infty$ wheres $\boldsymbol{\Sigma}_{wr}$ has bounded eigenvalues, we have $\lim_{N\to\infty}(\boldsymbol{A} + \boldsymbol{I}_K)^{-1} = \boldsymbol{0}_{K\times K}$, hence $(II) \to \boldsymbol{\lambda}_{\tilde{v}}^\top$. □

## A.3 Estimating Time-Varying Risk Premia in Section 1.2

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 $\boldsymbol{\Sigma}_{\epsilon v} = \mathrm{cov}(\boldsymbol{\epsilon}_{vt}) = \boldsymbol{H}\boldsymbol{H}^\top$. We generalize the rotation invariance as follows:

$$\boldsymbol{r}_t = \boldsymbol{\alpha} - \frac{\boldsymbol{\Upsilon}_r}{2} + \underbrace{\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\boldsymbol{\lambda}_{\tilde{v}}}_{\boldsymbol{\lambda}_v} + \underbrace{\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\tilde{\boldsymbol{v}}_t}_{\boldsymbol{v}_t} + \boldsymbol{w}_{rt}, \quad g_t = \mu_g + \sum_{s=0}^{\bar{S}}\tilde{\rho}_s\underbrace{\tilde{\boldsymbol{\eta}}_g^\top\boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_g^\top}\underbrace{\boldsymbol{H}\boldsymbol{\epsilon}_{\tilde{v},t-s}}_{\boldsymbol{\epsilon}_{v,t-s}} + w_{gt},$$

$$m_t = \kappa_m - \boldsymbol{\lambda}_v^\top(\boldsymbol{H}^{-1})^\top\boldsymbol{H}^{-1}\boldsymbol{\epsilon}_{vt} - \boldsymbol{\mu}_{v,t-1}^\top(\boldsymbol{H}^{-1})^\top\boldsymbol{H}^{-1}\boldsymbol{\epsilon}_{vt} = \kappa_m - \boldsymbol{\lambda}_v^\top\boldsymbol{\Sigma}_{\epsilon v}^{-1}\boldsymbol{\epsilon}_{vt} - \boldsymbol{\mu}_{v,t-1}^\top\boldsymbol{\Sigma}_{\epsilon v}^{-1}\boldsymbol{\epsilon}_{vt}, \text{ and}$$

$$\lambda_{g,t-1}^S = \frac{\sum_{\tau=0}^S\sum_{s=0}^{\tau}\tilde{\rho}_s}{1+S}\cdot\underbrace{\tilde{\boldsymbol{\eta}}_g^\top\boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_g^\top}\underbrace{\boldsymbol{H}(\boldsymbol{\lambda}_{\tilde{v}} + \mathbb{E}_{t-1}\left[\boldsymbol{\mu}_{\tilde{v},t+\tau-s-1}\right])}_{\boldsymbol{\lambda}_v + \mathbb{E}_{t-1}\left[\boldsymbol{\mu}_{v,t+\tau-s-1}\right]}. \tag{A1}$$

Thus, impulse response coefficients $\{\tilde{\rho}_s\}$ and risk premia $\lambda_{g,t}^S$ are point identified.

**Proposition A1** (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 ($\sigma_{wg}^2$, $\boldsymbol{\rho}_g$, and $\boldsymbol{\eta}_g$) follow the normal-inverse-gamma distribution in equations (IA.1)–(IA.3) of Internet Appendix IA.1.2. 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.1.2.*

*(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 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 $0 \leq S \leq \bar{S}$.*

The proof of the proposition is reported in Internet Appendix IA.1.3.

## A.4 Additional Figures and Tables

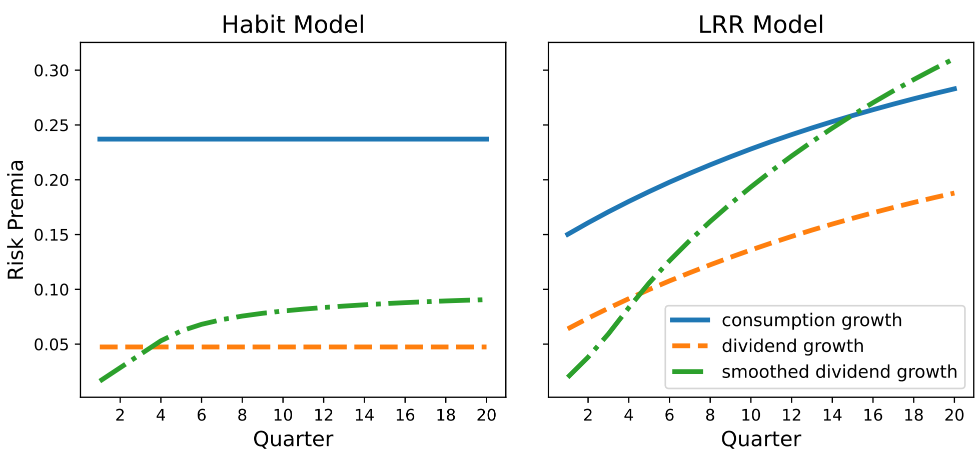



**Figure A1:** Term Structure of Risk Premia in Habit and Long-Run Risk Models

Term structure of risk premia in Campbell and Cochrane (1999) habit model (left panel) and Bansal and Yaron (2004) long-run risk model (right panel) for three variables: quarterly consumption growth, quarterly dividend growth, and quarterly growth in smoothed dividend payments (aggregate dividends paid in previous 12 months). Risk premia normalized by quarterly volatility. Calibration details in Internet Appendix IA.5.

**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. $\lambda_g^S$ is defined in equation (6), and $\lambda_g^{S,\star}$ is $\lambda_g^S$'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

Svetlana Bryzgalova,[a] Jiantao Huang,[b] and Christian Julliard[c*]

[a]*London Business School*

[b]*University of Hong Kong*

[c]*London School of Economics, FMG, SRC, and CEPR*

**Abstract**

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

[*]*Email addresses*: `sbryzgalova@london.edu` (S. Bryzgalova), `huangjt@hku.hk` (J. Huang), and `c.julliard@lse.ac.uk` (C. Julliard).

# IA.1 Additional Propositions and Proofs

## IA.1.1 Constructive derivation of Theorem 1

To build intuition, consider an explicit derivation of the representation in Theorem 1. By definition, the priced component $\tilde{g}_t$ satisfies $\tilde{g}_t \in \mathcal{M}_t = \overline{\text{span}}\{\tilde{m}_s : s \leq t\}$, while the unpriced component $\tilde{w}_t^g := g_t - \tilde{g}_t - \mu_g$ satisfies $\tilde{w}_t^g \perp \mathcal{M}_t$.

*i) Finite projections.* For each $n \geq 0$, let $\mathcal{M}_t^{(n)} = \text{span}\{\tilde{m}_t, \tilde{m}_{t-1}, \ldots, \tilde{m}_{t-n}\}$ and define the finite projection of the priced component: $\tilde{g}_t^{(n)} = \Pi_{\mathcal{M}_t^{(n)}}(\tilde{g}_t) = \sum_{j=0}^{n} \alpha_j^{(n)} \tilde{m}_{t-j}$, for some coefficients $\alpha_j^{(n)}$ determined by the least-squares projection.

*ii) Substitute the Wold representation of $\tilde{m}_t$.* By the Wold decomposition, $\tilde{m}_{t-j} = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-j-k}^m + V_{t-j}^m$, where $\varepsilon_t^m = \tilde{m}_t - \Pi_{\mathcal{M}_{t-1}}(\tilde{m}_t)$ are the innovations and $V_t^m$ is the deterministic component (perfectly predictable from the infinite past). Substituting into $\tilde{g}_t^{(n)}$ gives:

$$\tilde{g}_t^{(n)} = \sum_{j=0}^{n} \alpha_j^{(n)} \left( \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-j-k}^m + V_{t-j}^m \right).$$

*iii) Collect terms by lag order.* Setting $s = j + k$, rearrange to obtain:

$$\tilde{g}_t^{(n)} = \sum_{s=0}^{\infty} \left( \sum_{j=0}^{\min(n,s)} \alpha_j^{(n)} \psi_{s-j} \right) \varepsilon_{t-s}^m + \sum_{j=0}^{n} \alpha_j^{(n)} V_{t-j}^m.$$

Define $\theta_s^{(n)} := \sum_{j=0}^{\min(n,s)} \alpha_j^{(n)} \psi_{s-j}$, $W_t^{(n)} := \sum_{j=0}^{n} \alpha_j^{(n)} V_{t-j}^m$, so that $\tilde{g}_t^{(n)} = \sum_{s=0}^{\infty} \theta_s^{(n)} \varepsilon_{t-s}^m + W_t^{(n)}$.

*iv) Take limits.* Because $\tilde{g}_t \in \mathcal{M}_t$ and $\mathcal{M}_t^{(n)} \uparrow \mathcal{M}_t$, we have $\tilde{g}_t^{(n)} \xrightarrow{L^2} \tilde{g}_t$ as $n \to \infty$. By the continuity of the inner product in $L^2$, for each fixed $s$: $\theta_s^{(n)} = \frac{\mathbb{E}[\tilde{g}_t^{(n)} \varepsilon_{t-s}^m]}{\sigma_\varepsilon^2} \to \frac{\mathbb{E}[\tilde{g}_t \varepsilon_{t-s}^m]}{\sigma_\varepsilon^2} =: \theta_s$. Since each $W_t^{(n)} \in \mathcal{V}$ (the deterministic subspace), their limit $W_t$ also lies in $\mathcal{V}$. Hence, taking limits yields the priced component representation: $\tilde{g}_t = \sum_{s=0}^{\infty} \theta_s \varepsilon_{t-s}^m + W_t$. Adding back the mean and the unpriced component $\tilde{w}_t^g$, we obtain the full decomposition: $g_t = \mu_g + \sum_{s=0}^{\infty} \theta_s \varepsilon_{t-s}^m + w_t^g$, $w_t^g := W_t + \tilde{w}_t^g$, where $w_t^g$ is orthogonal to $m_t$, and $\theta_j = \mathbb{E}[(g_t - \mu_g)\, \varepsilon_{t-j}^m] / \sigma_\varepsilon^2$.

## IA.1.2 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:*

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

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

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

*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, defining $\boldsymbol{B}_r^\top = (\boldsymbol{\mu}_r, \boldsymbol{\beta}_v)$, we update model parameters in $\boldsymbol{r}_t$'s equation using a normal-inverse-Wishart distribution, as follows:*

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

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

*(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:*

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

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

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

*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:

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

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

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

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

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

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 $\sigma_{wg}^2$ is

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

hence, the posterior distribution of $\sigma_{wg}^2$ 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,

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

where $\hat{\boldsymbol{\rho}}_g = \left(\boldsymbol{V}_{\boldsymbol{\rho}}^\top\boldsymbol{V}_{\boldsymbol{\rho}}\right)^{-1}\boldsymbol{V}_{\boldsymbol{\rho}}^\top\boldsymbol{G}$. 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 $\sigma_{wg}^2\left(\boldsymbol{V}_{\boldsymbol{\rho}}^\top\boldsymbol{V}_{\boldsymbol{\rho}}\right)^{-1}$ with the Newey and West (1987) type of sandwich covariance matrix, denoted as $\hat{\boldsymbol{\Sigma}}_\rho$, as follows:

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

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

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

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

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

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

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

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 $\{\sigma_{wr,n}^2\}_{n=1}^{N}$.

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

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

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

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

which implies equation (IA.6). To avoid potential numerical difficulty in the Gibbs sampler, we replace $\left(\boldsymbol{\beta}_{\boldsymbol{v}}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{wr}}^{-1}\boldsymbol{\beta}_{\boldsymbol{v}}\right)^{-1}$ with $\left(\boldsymbol{\beta}_{\boldsymbol{v}}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{wr}}^{-1}\boldsymbol{\beta}_{\boldsymbol{v}}+\boldsymbol{I}_K\right)^{-1}$ when sampling $\boldsymbol{v}_t$ in equation (IA.6). The posterior distribution of $(\boldsymbol{\mu}_v, \boldsymbol{\Sigma}_v)$ is

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

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

### IA.1.3 Proof of Proposition A1

The only new ingredient in Proposition A1 is step 4, which estimates the model parameters in the VAR(q) system of $\boldsymbol{x}_t$. First, we introduce the following matrix notations:

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

and equation (22) implies that the data likelihood is

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

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,

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

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

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

## IA.2 Macro ARMA Processes: Calibration and Selection

We consider a general simulation setup for $g_t$ as follows:

$$g_t = \mu_g + x_{t-1} + \sigma_{t-1}e_{gt},\ x_t = \rho_x x_{t-1} + \varphi_x \sigma_{t-1}e_{xt},\ \sigma_t^2 = \sigma^2(1-\nu) + \nu\sigma_{t-1}^2 + \sigma_\sigma e_{\sigma t}, \tag{IA.11}$$

where $e_{gt}, e_{xt}, e_{\sigma t} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$. In particular, $x_{t-1}$ is the slow-moving component in $g_t$; $e_{xt}$ is the shock to the persistent component; $e_{gt}$ is the contemporaneous shock to $g_t$; $e_{\sigma t}$ is the shock to the potentially time-varying volatility. Then, $g_t$ can be calibrated to match several model specifications in the literature:

**Table IA.I:** ARIMA model selection (240 quarters)

| Bansal et al. (2016): consumption | | | | Croce (2014): TFP | | | | Belo and Li (2023): profitability | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| BIC | | AIC | | BIC | | AIC | | BIC | | AIC | |
| $p\ d\ q$ | freq | $p\ d\ q$ | freq | $p\ d\ q$ | freq | $p\ d\ q$ | freq | $p\ d\ q$ | freq | $p\ d\ q$ | freq |
| 0 0 0 | 78.1% | 0 0 0 | 51.0% | 0 0 0 | 60.7% | 0 0 0 | 32.6% | 1 0 1 | 31.6% | 1 0 1 | 30.3% |
| 0 1 1 | 8.0% | 0 0 1 | 6.3% | 0 1 1 | 16.5% | 1 0 1 | 11.2% | 0 1 1 | 25.6% | 0 1 1 | 12.3% |
| 1 1 1 | 3.0% | 1 0 1 | 5.7% | 1 1 1 | 6.0% | 0 1 1 | 8.5% | 0 0 0 | 13.3% | 1 1 1 | 6.2% |
| 1 0 0 | 2.9% | 1 0 0 | 5.2% | 1 0 1 | 4.4% | 0 0 1 | 6.6% | 1 0 0 | 9.7% | 2 0 1 | 6.1% |
| 0 0 1 | 1.7% | 2 0 2 | 4.4% | 1 0 0 | 3.3% | 1 1 1 | 4.8% | 2 0 0 | 6.5% | 1 0 2 | 4.8% |
| 1 1 2 | 1.2% | 0 1 1 | 3.4% | 2 1 1 | 2.3% | 1 0 0 | 4.0% | 1 1 1 | 4.8% | 2 0 0 | 3.8% |
| 2 1 1 | 1.2% | 1 1 1 | 2.9% | 0 0 1 | 1.9% | 1 1 2 | 3.5% | 3 0 0 | 2.3% | 2 0 2 | 3.5% |
| 1 0 1 | 0.6% | 1 1 2 | 2.1% | 1 1 2 | 1.6% | 2 0 2 | 2.9% | 0 0 1 | 1.6% | 0 1 2 | 3.4% |
| 3 1 1 | 0.6% | 2 0 0 | 2.1% | 3 1 1 | 0.5% | 2 0 0 | 2.5% | 0 1 2 | 1.0% | 1 1 2 | 3.3% |
| 2 0 0 | 0.4% | 1 0 2 | 1.3% | 2 1 2 | 0.4% | 2 0 1 | 2.5% | 1 1 2 | 0.9% | 0 0 1 | 3.1% |

Frequency of ARIMA(p,d,q) models selected by BIC and AIC in 1,000 simulations of the latent AR(1) model in equation (IA.11). Top 10 most frequently selected models shown.

(1) Log consumption growth in Bansal and Yaron (2004). We calibrate the quarterly log growth following Table 6, "Quarterly" column, of Bansal, Kiku, and Yaron (2016).

(2) Log productivity/TFP growth in Croce (2014). We adopt the quarterly calibration therein, which studies a constant volatility process in the baseline analysis.

(3) Log aggregate profitability growth in Belo and Li (2023). We calibrate the quarterly log profitability growth using the parameter estimates in Table 1 therein.

BIC and AIC specification selection results for the above specifications are reported in Table IA.I. Overwhelmingly, in samples as long as the historical ones, canonical specification selection fails to recover the pseudo-true calibrated processes with very high probability, and very often, the selected specification implies no predictability in the macro quantities.

## IA.3 Simulations: Time-Varying Risk Premia

In this section, we explore the finite-sample performance of our Bayesian estimator in Proposition A1 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, $\tilde{\boldsymbol{v}}_t = \hat{\boldsymbol{\phi}}_1 \tilde{\boldsymbol{v}}_{t-1} + \boldsymbol{\epsilon}_{\tilde{v}t}$, $\boldsymbol{\epsilon}_{\tilde{v}t} \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}_K, \boldsymbol{I}_K)$, where $\hat{\boldsymbol{\phi}}_1$ is calibrated by running the VAR(1) regression using the top five PCs of asset returns.[2] 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 $\tilde{\boldsymbol{v}}_t$: $g_t = c \cdot \sum_{s=0}^{\bar{S}} \hat{\rho}_s f_{t-s} + w_{gt}$, $f_t = \frac{1}{\sqrt{3}}(1,0,1,0,1)\boldsymbol{\epsilon}_{\tilde{v}t}$.

Similar to the simulations in the main text, we report the size, power, and time series fit

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

in $g_t$'s equation ($R_g^2$) and the correlation between estimated and pseudo-true latent process $f_t$ in Tables IA.VIII–IA.IX. Overall, our Bayesian estimator in Proposition A1 has satisfactory finite-sample performance, delivering consistent estimates of unconditional risk premia.

## IA.4 Data Description

We consider a cross-section of 275 equity portfolios collected from Ken French's website (FF275): 25 ($5 \times 5$) 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 3. 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.

## IA.5 Term Structure of Risk Premia in Structural 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: log SDF $= m_{t+1} = \log\delta - \gamma g + \gamma(1-\phi)(s_t - \bar{s}) - \gamma[1+\lambda(s_t)]v_{t+1}$; log consumption surplus ratio $= s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)v_{t+1}$; log consumption growth $= \Delta c_{t+1} = g + v_{t+1}$, $v_{t+1} \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma^2)$; log dividend growth $= \Delta d_{t+1} = g + w_{t+1}$, $w_{t+1} \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma_w^2)$, with $\text{corr}(w_t, v_t) = \rho$.

CC choose the specification of $\lambda(s_t)$ to ensure a constant risk-free rate:

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

Second, 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. The dynamics of the state variables are: conditional consumption mean $= x_{t+1} = \rho x_t + \varphi_e \sigma_t e_{t+1}$; log consumption growth $= \Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{t+1}$; log dividend growth $= \Delta d_{t+1} = \mu_d + \phi_d x_t + \pi\sigma_t\eta_{t+1} + \varphi_d\sigma_t u_{t+1}$; stochastic volatility $= \sigma_{t+1}^2 = \sigma^2 + \nu_1(\sigma_t^2 - \sigma^2) + \sigma_\omega \omega_{t+1}$, 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 log price-consumption ratio, obtaining the log SDF

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

where $\lambda$s are analytical functions of the preference parameters and approximation constants.

We simulate the dynamics of all quantities of interest following exactly the same calibrations as in Table 1 of CC for their model, and as in Bansal et al. (2012) for the BY model. We simulate the monthly sequences and aggregate them into quarterly observations. Using the quarterly data, we calculate the unconditional risk premia of consumption and dividend growth as follows:

$$\lambda_g^S = -\frac{\mathbb{E}\big[\mathrm{cov}_t\big(\bar{m}_{t\to t+S},\ g_{t\to t+S}\big)\big]}{S\cdot\sigma(g_{t+1})}, \tag{IA.12}$$

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

Using equation (IA.12), we can obtain closed-form solutions for the term structure of consumption and dividend risk premia in the external habit model: $\lambda_{\Delta c}^S = \gamma\sigma\big[1 + \mathbb{E}[\lambda(s_t)]\big]$, $\lambda_{\Delta d}^S = \rho\gamma\sigma\big[1 + \mathbb{E}[\lambda(s_t)]\big]$.[3] 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 $\lambda_{\Delta c}^S$ and $\lambda_{\Delta d}^S$. In contrast, we do not have simple closed-form solutions in the LRR model and 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.12)? 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.12) 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, $R_{f,t\to t+S} =$

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

$\prod_{\tau=1}^{S} R_{f,t+\tau-1\to t+\tau}$ 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 $R_{t\to t+S}$,

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

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 $\mathbb{E}\big[\tilde{M}_{t+\tau-1\to t+\tau}\big] = \mathbb{E}_{t+\tau-1}\big[\tilde{M}_{t+\tau-1\to t+\tau}\big] = 1$. From the fundamental asset pricing equation we have:

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

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.

Assuming that asset returns, macro variables, and the SDF follow log-normal distributions and denoting $\prod_{\tau=1}^{S}\tilde{M}_{t+\tau-1\to t+\tau} = \exp\left\{\sum_{\tau=1}^{S}\tilde{m}_{t+\tau-1\to t+\tau}\right\} = \exp\{\tilde{m}_{t\to t+S}\}$ and $\prod_{\tau=1}^{S}\frac{R_{t+\tau-1\to t+\tau}}{R_{f,t+\tau-1\to t+\tau}} = \exp\{\tilde{r}^e_{t\to t+S}\}$ we have

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

Removing the common component, $\exp\{\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big) + \frac{1}{2}\mathrm{var}_t\big(\tilde{r}^e_{t\to t+S}\big)\}$, from both the left and right sides of equation (IA.13) and rearranging terms we have: $\mathbb{E}_t\big(\tilde{r}^e_{t\to t+S}\big) + \frac{1}{2}\mathrm{var}_t\big(\tilde{r}^e_{t\to t+S}\big) = -\mathrm{cov}_t\big(\tilde{m}_{t\to t+S}, \tilde{r}^e_{t\to t+S}\big)$. Therefore, $-\mathrm{cov}_t\big(\tilde{m}_{t\to t+S}, \tilde{r}^e_{t\to t+S}\big)$ 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: $\tilde{m}_{t+1} = m_{t+1} - \log(\mathbb{E}_t[M_{t+1}]) = m_{t+1} - \mathbb{E}_t[m_{t+1}] - \frac{1}{2}\mathrm{var}_t[m_{t+1}] = \bar{m}_{t+1} - \frac{1}{2}\mathrm{var}_t[m_{t+1}]$. It is easy to show that $\mathrm{var}_t[m_{t+1}]$ does not correlate with consumption or dividend growth in the habit and long-run risk formulations considered. Finally, dividing the multi-period risk premia, $-\mathrm{cov}_t\big(\tilde{m}_{t\to t+S}, \tilde{r}^e_{t\to t+S}\big)$, by the number of periods $S$, and normalizing by the volatility of the single-period variable, leads to the definition in equation (IA.12).

## IA.6 Alternative Formulation for $g_t$

The priced Wold decomposition in Theorem 1 implies that the factor $g_t$ can be driven by both $\boldsymbol{\epsilon}_{\tilde{\boldsymbol{v}}\boldsymbol{t}}$ and $\boldsymbol{\mu}_{\tilde{v},t-1}^{\top}\boldsymbol{\epsilon}_{\tilde{v},t}$, where the latter captures time-varying risk premia of latent factors. In this appendix, we rewrite the dynamics of $g_t$ as follows:

$$g_t = \mu_g + \sum_{s=0}^{\bar{S}} \tilde{\rho}_s \big(\tilde{\boldsymbol{\eta}}_{g1}^{\top}\boldsymbol{\epsilon}_{\tilde{v},t-s} + \tilde{\eta}_{g2}\boldsymbol{\mu}_{\tilde{v},t-s-1}^{\top}\boldsymbol{\epsilon}_{\tilde{v},t-s}\big) + w_{gt} = \mu_g + \sum_{s=0}^{\bar{S}} \tilde{\rho}_s \underbrace{\tilde{\boldsymbol{\eta}}_g^{\top}\boldsymbol{\epsilon}_{g,t-s}}_{f_{t-s}} + w_{gt}, \tag{IA.14}$$

where $\tilde{\boldsymbol{\eta}}_g^{\top}\tilde{\boldsymbol{\eta}}_g = 1$, $\tilde{\boldsymbol{\eta}}_g^{\top} = (\tilde{\boldsymbol{\eta}}_{g1}^{\top}, \tilde{\eta}_{g2})^{\top}$ and $\boldsymbol{\epsilon}_{g,t-s}^{\top} = (\boldsymbol{\epsilon}_{\tilde{v},t-s}^{\top}, \boldsymbol{\mu}_{\tilde{v},t-s-1}^{\top}\boldsymbol{\epsilon}_{\tilde{v},t-s})^{\top}$. Under this assumption, we show that the term structure of $g$'s risk premia is

$$\lambda_{g,t-1}^{S} = \sum_{\tau=0}^{S}\sum_{s=0}^{\tau} \tilde{\rho}_s \frac{\tilde{\boldsymbol{\eta}}_{g1}^{\top}\boldsymbol{\lambda}_{\tilde{v}} + (\tilde{\boldsymbol{\eta}}_{g1} + \tilde{\eta}_{g2}\boldsymbol{\lambda}_{\tilde{v}})^{\top}\mathbb{E}_{t-1}\big[\boldsymbol{\mu}_{\tilde{v},t+\tau-s-1}\big] + \tilde{\eta}_{g2}\mathbb{E}_{t-1}\big[\boldsymbol{\mu}_{\tilde{v},t+\tau-s-1}^{\top}\boldsymbol{\mu}_{\tilde{v},t+\tau-s-1}\big]}{1+S}. \tag{IA.15}$$

If $\tilde{\eta}_{g2} = 0$, the above definition is exactly identical to that in equation (21).

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 $\boldsymbol{\Sigma}_{\epsilon v} = \mathrm{cov}(\boldsymbol{\epsilon}_{vt}) = \boldsymbol{H}\boldsymbol{H}^{\top}$. We generalize the rotation invariance to identify the time-varying risk premia as follows:

$$\begin{aligned}
\boldsymbol{r}_t &= \boldsymbol{\alpha} - \frac{\boldsymbol{\Upsilon}_r}{2} + \underbrace{\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\boldsymbol{\lambda}_{\tilde{v}}}_{\boldsymbol{\lambda}_v} + \underbrace{\boldsymbol{\beta}_{\tilde{v}}\boldsymbol{H}^{-1}}_{\boldsymbol{\beta}_v}\underbrace{\boldsymbol{H}\tilde{\boldsymbol{v}}_t}_{\boldsymbol{v}_t} + \boldsymbol{w}_{rt},\\
m_t &= \kappa_m - \boldsymbol{\lambda}_v^{\top}(\boldsymbol{H}^{-1})^{\top}\boldsymbol{H}^{-1}\boldsymbol{\epsilon}_{vt} - \boldsymbol{\mu}_{v,t-1}^{\top}(\boldsymbol{H}^{-1})^{\top}\boldsymbol{H}^{-1}\boldsymbol{\epsilon}_{vt} = \kappa_m - \boldsymbol{\lambda}_v^{\top}\boldsymbol{\Sigma}_{\epsilon v}^{-1}\boldsymbol{\epsilon}_{vt} - \boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\epsilon v}^{-1}\boldsymbol{\epsilon}_{vt},\\
g_t &= \mu_g + \sum_{s=0}^{\bar{S}} \tilde{\rho}_s \Bigg(\underbrace{\tilde{\boldsymbol{\eta}}_{g1}^{\top}\boldsymbol{H}^{-1}}_{\boldsymbol{\eta}_{g1}^{\top}}\underbrace{\boldsymbol{H}\boldsymbol{\epsilon}_{\tilde{v},t-s}}_{\boldsymbol{\epsilon}_{v,t-s}} + \eta_{g2}\underbrace{\boldsymbol{\mu}_{v,t-s-1}^{\top}(\boldsymbol{H}^{-1})^{\top}}_{\boldsymbol{\mu}_{\tilde{v},t-s-1}^{\top}}\underbrace{\boldsymbol{H}^{-1}\boldsymbol{\epsilon}_{v,t-s}}_{\boldsymbol{\epsilon}_{\tilde{v},t-s}}\Bigg) + w_{gt}, \text{ and}\\
\lambda_{g,t-1}^{S} &= \sum_{\tau=0}^{S}\sum_{s=0}^{\tau} \tilde{\rho}_s \frac{\boldsymbol{\eta}_{g1}^{\top}\boldsymbol{\lambda}_v + (\boldsymbol{\eta}_{g1} + \tilde{\eta}_{g2}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}}^{-1}\boldsymbol{\lambda}_v)^{\top}\mathbb{E}_{t-1}\big[\boldsymbol{\mu}_{v,t+\tau-s-1}\big] + \tilde{\eta}_{g2}\mathbb{E}_{t-1}\big[\boldsymbol{\mu}_{v,t+\tau-s-1}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}}^{-1}\boldsymbol{\mu}_{v,t+\tau-s-1}\big]}{1+S}.
\end{aligned} \tag{IA.16}$$

We perform a formal test to determine whether the additional component, $\boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\epsilon v}^{-1}\boldsymbol{\epsilon}_{vt}$, is essential in driving the dynamics of $g_t$. First, Table IA.XI shows that including $\boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\epsilon v}^{-1}\boldsymbol{\epsilon}_{vt}$ in the dynamics of $g_t$ only marginally improves the time-series fit. Furthermore, we investigate the cumulative impulse response functions (CIRFs) of $g_t$ to one-standard-deviation innovations in $\boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}}\boldsymbol{\epsilon}_{v,t}$, that is, $\big(\sum_{s=0}^{\bar{S}}\rho_s\big)\cdot\eta_{g2}/\sigma\big(\boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}}\boldsymbol{\epsilon}_{v,t}\big)$. Except for HKM and VIX factors, the

CIRFs of other factors are not significantly different from zero. Moreover, estimates of CIRFs tend to be rather volatile for quarterly macro factors, leading to noisy risk-premia estimates. Even for HKM and VIX factors, the incremental $R_g^2$ is almost negligible—less than 1%—with the inclusion of $\boldsymbol{\mu}_{v,t-1}^\top \boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}} \boldsymbol{\epsilon}_{v,t}$. Therefore, omitting $\boldsymbol{\mu}_{v,t-1}^\top \boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}} \boldsymbol{\epsilon}_{v,t}$ in the $g_t$ dynamics, from the empirical perspective, results in at most a slight bias in risk premia estimates.

## IA.7 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 $t+s$. We define: holding period return $= R_{t,t+S} = \frac{D_{t+s}}{P_{s,t}}$; spot equity yield $= e_{s,t} = \frac{1}{s}\log\left(\frac{D_t}{P_{s,t}}\right)$; forward equity yield $= e_{s,t}^f = \frac{1}{s}\log\left(\frac{D_t}{P_{s,t}}\right) - \frac{1}{s} r_{f,t,t+s}$, where $\frac{1}{s} r_{f,t,t+s}$ is log risk-free rate with time-to-maturity $s$.

The fundamental asset pricing equation implies that $P_{s,t} = \mathbb{E}_t[M_{t,t+s} \cdot D_{t+s}]$, 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 $R_{t,t+s} = \frac{D_{t+s}}{P_{s,t}} = \frac{D_t}{P_{s,t}} \frac{D_{t+s}}{D_t}$, which implies the following per-period log strip return:

$$r_{t+s} = \frac{1}{s}\log(R_{t,t+s}) = \frac{1}{s}\log\left(\frac{D_t}{P_{s,t}}\right) + \frac{1}{s}\log\left(\frac{D_{t+s}}{D_t}\right) = e_{s,t} + g_{d,t,t+s}, \tag{IA.17}$$

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.17) also implies that the conditional variance of $r_{t+s}$ is identical to that of $g_{d,t,t+s}$, that is, $\mathrm{var}_t(r_{t+s}) = \mathrm{var}_t(g_{d,t,t+s})$.

Assuming the joint log normality of SDF and dividend growth and the SDF, we can rewrite:

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

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

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

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

$$\mathbb{E}_t[r_{t+s}] = e_{s,t} + \mathbb{E}_t[g_{d,t,t+s}] = \frac{1}{s} r_{f,t,t+s} - \frac{1}{2s}\mathrm{var}_t(\Delta d_{t,t+s}) - \frac{1}{s}\mathrm{cov}_t(m_{t,t+s}, \Delta d_{t,t+s}). \tag{IA.18}$$

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: $\mathbb{E}_t[r_{t+s}] + \frac{s}{2}\text{var}_t(r_{t+s}) = \frac{1}{s}r_{f,t,t+s} + \lambda^s_{dt}$. 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, $\lambda^s_{dt}$, 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.18), we can derive the forward equity yield, as: $e^f_{s,t} = e_{s,t} - \frac{1}{s}r_{f,t,t+s} = \lambda^s_{dt} - \mathbb{E}_t[g_{d,t,t+s}] - \frac{1}{2s}\text{var}_t(\Delta d_{t,t+s})$, where the last term, $\frac{1}{2s}\text{var}_t(\Delta d_{t,t+s})$ is negligible in the data.

# IA.8 Additional Tables and Figures

**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 |
| Total factor productivity (TFP) growth | Q3 1963 – Q4 2019 | John Fernald's website |
| Utilization-adjusted TFP growth in Fernald (2014) | Q3 1963 – Q4 2019 | John Fernald's website |
| Civilian unemployment rate (UNRATE) change | Q3 1963 – Q4 2019 | Federal Reserve Bank of St. Louis |
| Hours worked, log growth | Q3 1963 – Q4 2019 | Federal Reserve Bank of St. Louis |
| Real Investment per capita, log growth | Q3 1963 – Q4 2019 | Federal Reserve Bank of St. Louis |
| 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 liquidity factors (Pástor and Stambaugh (2003)) | Jul 1963 – Dec 2019 | Lubos Pastor's website |
| $\Delta\log(\text{VIX}_t) = \log(\text{VIX}_t) - \log(\text{VIX}_{t-1})$ | 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) | Jul 1963 – Dec 2019 | Ken French's website |

List of factors used in Section 3 with descriptions, samples, and data sources. Dividends: monthly real S&P500 dividend payments from Robert Shiller's website; to avoid mechanical seasonality, we use smoothed dividends $D_t$ (sum of previous 12 months) with growth rate $\log(D_t/D_{t-1})$. Following Angeletos, Collard, and Dellas (2020), hours worked = log(Nonfarm business sector: average weekly hours × Employment Level / Civilian non-institutional population); real investment per capita = log[(Share of GDP: personal durable consumption expenditures + Share of GDP: gross private domestic investment) × Real GDP per capita]. Term spread: difference between 10-year and 3-month Treasury yields. Default spread: difference between BAA and AAA corporate bond yields. Value spread: 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 |

Frequency of rejecting $H_0 : \lambda_g^S = \lambda_g^{S,\star}$ using 90%, 95%, and 99% credible intervals from Proposition 1, where $\lambda_g^S$ is defined in equation (6) and $\lambda_g^{S,\star}$ is the pseudo-true value. Strong factors with $R_g^2 \in \{10\%, 20\%, 30\%\}$. Simulations: monthly data, true model has 5 factors with $\bar{S} = 16$ and $T = 600$; estimated models use 4, 5, or 7 factors with $\bar{S} = 24$. 1,000 Monte Carlo simulations.

**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 |

Frequency of rejecting $H_0 : \lambda_g^S = 0$ using 90%, 95%, and 99% credible intervals from Proposition 1, where $\lambda_g^S$ is defined in equation (6). Useless factors with varying persistence: the persistent component in $g_t$ accounts for 10%, 20%, or 30% of time series variation. Simulations: quarterly data, true model has 5 factors with $T = 200$ and $g_t \perp \boldsymbol{r}_t$; estimated models use 4, 5, or 7 factors with $\bar{S} = 12$. 1,000 Monte Carlo simulations.

**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 |

Frequency of rejecting $H_0 : \lambda_g^S = 0$ using 90%, 95%, and 99% credible intervals from Proposition 1, where $\lambda_g^S$ is defined in equation (6). Useless factors with varying persistence: the persistent component in $g_t$ accounts for 10%, 20%, or 30% of time series variation. Simulations: monthly data, true model has 5 factors with $T = 600$ and $g_t \perp \boldsymbol{r}_t$; estimated models use 4, 5, or 7 factors with $\bar{S} = 24$. 1,000 Monte Carlo simulations.

**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 |

Bayesian estimates of $R_g^2$ and $\text{corr}(\hat{f}_t, f_t)$ for strong and useless factors, where $R_g^2$ measures the percentage of $g_t$'s time series variation explained by latent factors and $\text{corr}(\hat{f}_t, f_t)$ is the correlation between true $f_t$ and its estimate $\hat{f}_t = \hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{v}}_t$. For useless factors, only $R_g^2$ is reported. Each cell shows median, 5th, and 95th percentiles from 1,000 simulations. Strong and useless factors with varying persistence: the persistent component in $g_t$ accounts for 10%, 20%, or 30% of time series variation. Simulations: true model has 5 factors; estimated models use $K \in \{4, 5, 7\}$ factors with $\bar{S} = 12$ for $T = 200$ or $\bar{S} = 24$ 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: frequency of rejecting $H_0 : \lambda_g = \lambda_g^\star$ using 90%, 95%, and 99% credible intervals from our Bayesian estimates (Proposition 1) and frequentist test statistics (Giglio and Xiu (2021), Theorem 1), where $\lambda_g^\star$ is the pseudo-true value. Panel B: frequency of rejecting $H_0 : \lambda_g = 0$. 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 the contemporaneous $\tilde{\boldsymbol{v}}_t$ ($\bar{S} = 0$). We estimate several model configurations with different numbers of factors (5, 7). 1,000 Monte Carlo simulations.

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

| | **Panel A.** $R_g^2$ | | | | | | **Panel B.** $\mathrm{corr}(\hat{f}_t, f_t)$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 |

Bayesian estimates of $R_g^2$ and $\mathrm{corr}(\hat{f}_t, f_t)$ for strong factors, where $R_g^2$ measures the percentage of $g_t$'s time series variation explained by $\boldsymbol{\epsilon}_{vt}$ and $\mathrm{corr}(\hat{f}_t, f_t)$ is the correlation between true $f_t$ and its estimate $\hat{f}_t = \hat{\boldsymbol{\eta}}_g^\top \hat{\boldsymbol{\epsilon}}_{vt}$. Each cell shows median, 5th, and 95th percentiles from 1,000 simulations. Varying persistence: the persistent component in $g_t$ accounts for 10%, 20%, or 30% of time series variation. We simulate monthly or quarterly observations of $g_t$ and $\boldsymbol{r}_t$ assuming: true model has 5 factors; estimated models use $K \in \{5, 7\}$ factors with $\bar{S} = 12$ for $T = 200$ or $\bar{S} = 24$ for $T = 600$.

**Table IA.IX:** Testing unconditional risk premia of strong factors when factors command time-varying risk premia in simulations

| | $S=0$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Quarterly frequency ($T=200$)** | | | | | | | | | | | | | |
| **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 |
| | $S=0$ | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 | 22 | 24 |
| **Monthly frequency ($T=600$)** | | | | | | | | | | | | | |
| **Panel D:** $R_g^2 = 30\%$ | | | | | | | | | | | | | |
| 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 E:** $R_g^2 = 20\%$ | | | | | | | | | | | | | |
| 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 F:** $R_g^2 = 10\%$ | | | | | | | | | | | | | |
| 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 |

Frequency of rejecting $H_0: \lambda_g^S = \lambda_g^{S,\star}$ (unconditional risk premia: $\lambda_g^S = \sum_{\tau=0}^{S}\sum_{s=0}^{\tau}\frac{\rho_s \boldsymbol{\eta}_g^\top \boldsymbol{\lambda}_v}{1+S}$) using 90%, 95%, and 99% credible intervals from Proposition A1, where $\lambda_g^{S,\star}$ is the pseudo-true value. Strong factors with $R_g^2 \in \{10\%, 20\%, 30\%\}$. Simulations: true model has 5 VAR(1) factors with $\bar{S} = 8$ and $T = 200$ (quarterly, Panels A–C) or $\bar{S} = 16$ and $T = 600$ (monthly, Panels D–F); estimated models use 5 or 7 factors with $\bar{S} = 12$ (quarterly) or $\bar{S} = 24$ (monthly). 1,000 Monte Carlo simulations.

**Table IA.X:** Factors' risk premia: $\bar{S} = 0$, Bayesian and Frequentist estimates

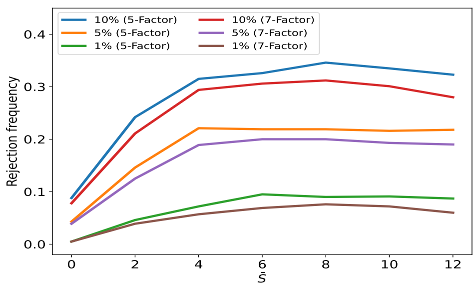
| | **Bayesian estimates** | | | | | | **Frequentist estimates (Giglio and Xiu, 2021)** | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\mathbb{E}[\lambda_g \mid \mathcal{D}]$ | | | $\mathbb{E}[R^2_g \mid \mathcal{D}]$ | | | $\mathbb{E}[\lambda_g \mid \mathcal{D}]$ | | | $\mathbb{E}[R^2_g \mid \mathcal{D}]$ | | |
| Number of factors: | 5 | 6 | 7 | 5 | 6 | 7 | 5 | 6 | 7 | 5 | 6 | 7 |
| **Panel A.** Original factors | | | | | | | | | | | | |
| Quarterly variables | | | | | | | | | | | | |
| AEM intermediary | 0.141*** | 0.176*** | 0.175*** | 10.4% | 12.2% | 12.4% | 0.160*** | 0.183*** | 0.184*** | 11.7% | 13.3% | 13.5% |
| Capital share growth | 0.032 | 0.013 | 0.013 | 1.8% | 2.8% | 2.8% | 0.031 | 0.018 | 0.015 | 2.0% | 2.4% | 3.4% |
| GDP growth | 0.005 | 0.013 | 0.013 | 4.2% | 4.3% | 4.3% | 0.007 | 0.010 | 0.009 | 4.2% | 4.2% | 4.3% |
| IP growth | -0.028 | 0.004 | 0.003 | 2.9% | 4.3% | 4.3% | -0.015 | -0.001 | -0.003 | 3.2% | 4.0% | 4.3% |
| Durable cons. growth | -0.011 | 0.000 | -0.002 | 7.7% | 7.9% | 8.2% | -0.015 | -0.002 | -0.006 | 7.1% | 7.7% | 8.7% |
| Nondurable cons. growth | 0.042 | 0.058 | 0.056 | 3.7% | 4.1% | 4.1% | 0.045 | 0.058 | 0.057 | 3.8% | 4.4% | 4.5% |
| Service cons. growth | 0.015 | 0.053 | 0.052 | 4.0% | 6.4% | 6.5% | 0.036 | 0.048 | 0.047 | 5.0% | 5.5% | 5.7% |
| Nondurable + service | 0.032 | 0.067* | 0.066* | 4.0% | 5.9% | 6.0% | 0.049 | 0.063* | 0.062* | 4.9% | 5.6% | 5.8% |
| Labor income growth | -0.006 | 0.035 | 0.028 | 1.5% | 3.8% | 8.7% | 0.016 | 0.027 | 0.020 | 2.8% | 3.3% | 7.9% |
| Dividend growth of SP500 | 0.037 | 0.037 | 0.044 | 5.1% | 5.4% | 11.7% | 0.045 | 0.035 | 0.043 | 4.3% | 4.7% | 10.8% |
| TFP growth | 0.014 | 0.002 | -0.003 | 6.9% | 7.3% | 8.2% | 0.003 | -0.001 | -0.006 | 7.1% | 7.2% | 9.8% |
| TFP growth (util.-adj.) | -0.016 | -0.048 | -0.054 | 2.9% | 4.9% | 6.7% | -0.037 | -0.047 | -0.052 | 4.4% | 4.7% | 7.4% |
| Unemployment rate chg. | -0.034 | -0.040 | -0.045 | 2.5% | 2.7% | 3.5% | -0.039 | -0.033 | -0.035 | 2.3% | 2.4% | 2.8% |
| Hours worked growth | 0.049 | 0.070* | 0.072* | 2.3% | 3.4% | 4.3% | 0.062* | 0.064* | 0.067* | 3.0% | 3.0% | 3.8% |
| Investment growth | 0.000 | 0.014 | 0.015 | 2.2% | 2.5% | 3.1% | 0.003 | 0.009 | 0.011 | 2.1% | 2.2% | 2.5% |
| Monthly variables | | | | | | | | | | | | |
| Oil price change | -0.018 | -0.017 | -0.016 | 2.6% | 4.5% | 4.5% | -0.015 | -0.017 | -0.004 | 3.0% | 3.6% | 5.4% |
| TED spread change | -0.034 | -0.040* | -0.033 | 6.8% | 10.9% | 17.4% | -0.042** | -0.044** | -0.055*** | 6.4% | 7.0% | 8.7% |
| Nontraded HKM interm. | 0.100*** | 0.104*** | 0.104*** | 60.3% | 61.0% | 61.1% | 0.099*** | 0.101*** | 0.104*** | 61.3% | 61.8% | 61.8% |
| Traded HKM interm. | 0.112*** | 0.116*** | 0.116*** | 70.3% | 71.0% | 71.2% | 0.110*** | 0.112*** | 0.114*** | 71.5% | 72.1% | 72.1% |
| PS liquidity | 0.061*** | 0.059*** | 0.062*** | 11.9% | 12.2% | 12.9% | 0.062*** | 0.060*** | 0.053*** | 11.9% | 12.3% | 12.4% |
| $\Delta \log(\text{VIX})$ | -0.120*** | -0.118*** | -0.118*** | 42.8% | 43.0% | 43.1% | -0.118*** | -0.118*** | -0.118*** | 42.8% | 42.9% | 42.9% |
| **Panel B.** AR(1) shocks of quarterly macro factors | | | | | | | | | | | | |
| AEM intermediary | 0.144*** | 0.182*** | 0.182*** | 10.8% | 12.9% | 13.3% | 0.166*** | 0.189*** | 0.190*** | 12.3% | 14.0% | 14.3% |
| Capital share growth | 0.041 | 0.017 | 0.017 | 2.1% | 3.5% | 3.7% | 0.039 | 0.019 | 0.016 | 2.0% | 3.0% | 4.0% |
| GDP growth | 0.005 | 0.012 | 0.012 | 4.2% | 4.3% | 4.3% | 0.007 | 0.010 | 0.009 | 4.2% | 4.2% | 4.3% |
| IP growth | -0.021 | 0.001 | -0.002 | 3.6% | 4.3% | 4.9% | -0.009 | -0.003 | -0.008 | 3.8% | 3.9% | 6.1% |
| Durable cons. growth | -0.010 | 0.002 | 0.000 | 7.4% | 7.6% | 7.8% | -0.013 | 0.000 | -0.003 | 6.8% | 7.4% | 8.1% |
| Nondurable cons. growth | 0.042 | 0.058 | 0.057 | 3.7% | 4.1% | 4.1% | 0.045 | 0.058 | 0.057 | 3.8% | 4.4% | 4.4% |
| Service cons. growth | 0.016 | 0.055 | 0.055 | 3.7% | 6.3% | 6.4% | 0.037 | 0.051 | 0.050 | 4.9% | 5.5% | 5.6% |
| Nondurable + service | 0.031 | 0.069* | 0.069* | 3.7% | 6.0% | 6.0% | 0.049 | 0.065* | 0.065* | 4.7% | 5.6% | 5.6% |
| Labor income growth | -0.007 | 0.033 | 0.027 | 1.4% | 3.6% | 8.3% | 0.015 | 0.026 | 0.019 | 2.7% | 3.1% | 7.3% |
| Dividend growth of SP500 | 0.067* | 0.068* | 0.074* | 3.3% | 3.4% | 9.5% | 0.068** | 0.066* | 0.073* | 2.9% | 2.9% | 8.0% |
| TFP growth | 0.019 | 0.006 | 0.001 | 7.2% | 7.7% | 8.6% | 0.007 | 0.003 | -0.002 | 7.5% | 7.6% | 10.1% |
| TFP growth (util.-adj.) | -0.012 | -0.044 | -0.050 | 3.1% | 5.1% | 6.9% | -0.033 | -0.043 | -0.048 | 4.6% | 5.0% | 7.6% |
| Unemployment rate chg. | -0.038 | -0.041 | -0.045 | 2.7% | 2.8% | 3.3% | -0.042 | -0.034 | -0.035 | 2.3% | 2.5% | 2.7% |
| Hours worked growth | 0.055 | 0.074** | 0.075* | 2.6% | 3.7% | 4.2% | 0.067** | 0.068* | 0.070* | 3.3% | 3.3% | 3.7% |
| Investment growth | -0.005 | 0.007 | 0.006 | 2.2% | 2.5% | 2.5% | -0.001 | 0.003 | 0.003 | 2.1% | 2.2% | 2.2% |

Factors' risk premia and time series fit $R^2_g$ using Bayesian methods (left panel) and Giglio and Xiu (2021) frequentist method (right panel). Panel A: original variables; Panel B: AR(1) shocks of macro factors. Bayesian estimates from Proposition 1 with $\bar{S} = 0$. Base assets: 275 Fama-French characteristic-sorted portfolios. Five-, six-, and seven-factor models for returns. Significance: * (**, ***) indicates 90% (95%, 99%) credible interval excludes zero for Bayesian estimates or 10% (5%, 1%) significance for frequentist estimates. Data sources in Internet Appendix IA.4.

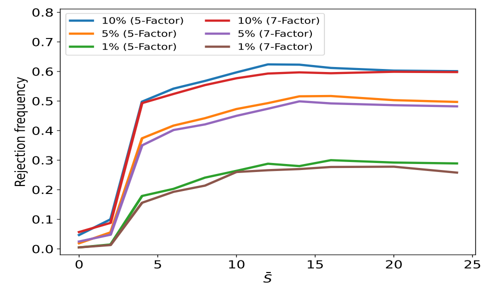
**Table IA.XI:** Time-series fit $R_g^2$ with and without $\boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}}\boldsymbol{\epsilon}_{v,t}$.

| | Without $\boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}}\boldsymbol{\epsilon}_{v,t}$ | With $\boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}}\boldsymbol{\epsilon}_{v,t}$ |
|---|---|---|
| AEM intermediary | 16.8% | 16.7% |
| Capital share growth | 9.5% | 10.0% |
| GDP growth | 21.4% | 23.4% |
| IP growth | 36.3% | 39.1% |
| Durable consumption growth | 18.7% | 19.6% |
| Nondurable consumption growth | 19.0% | 19.3% |
| Service consumption growth | 13.0% | 14.9% |
| Nondurable + service | 17.0% | 17.5% |
| Labor income growth | 8.0% | 8.1% |
| Dividend growth of S&P500 | 20.2% | 21.1% |
| TFP growth | 16.8% | 18.0% |
| TFP growth (utilization-adjusted) | 10.6% | 10.6% |
| Unemployment rate change | 30.9% | 33.8% |
| Hours worked growth | 25.4% | 27.8% |
| Investment growth | 28.1% | 29.5% |
| Oil price change | 9.2% | 9.4% |
| TED spread change | 10.7% | 10.7% |
| Nontraded HKM intermediary | 61.5% | 62.2% |
| Traded HKM intermediary | 71.4% | 72.2% |
| PS liquidity | 16.3% | 16.5% |
| $\Delta$ log (VIX) | 53.7% | 54.2% |

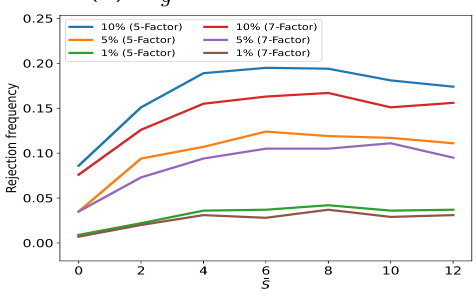
Time-series fit $R_g^2$ with and without $\boldsymbol{\mu}_{v,t-1}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{\epsilon v}}\boldsymbol{\epsilon}_{v,t}$ in the $g_t$ process. Column "Without": estimates from equation (20); column "With": estimates from equation (IA.14). Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns. For quarterly (monthly) factors, we use 12-quarter (24-month) lags in $g_t$ equations. Latent factors follow a VAR(1) process where both factors and four external predictors (PE ratio of S&P 500, term spread, default spread, and value spread) are driven by lagged external predictors. Data sources in Appendix IA.4.



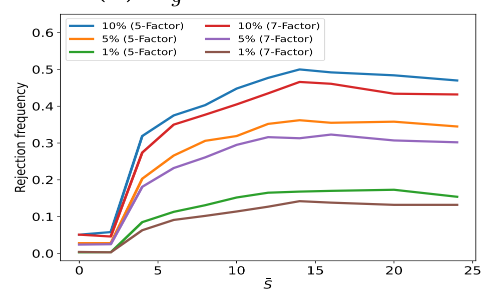

**(a)** $R_g^2 = 20\%$ & $T = 200$




**(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

Frequency of rejecting $H_0 : \lambda_g^{\bar{S}} = 0$ using 90%, 95%, and 99% credible intervals from Proposition 1, where $\lambda_g^{\bar{S}}$ is defined in equation (6). Strong factors with $R_g^2 \in \{10\%, 20\%, 30\%\}$ and sample sizes $T \in \{200, 600\}$. Each simulation estimates multiple model configurations varying the number of factors and $\bar{S}$. 1,000 simulations.

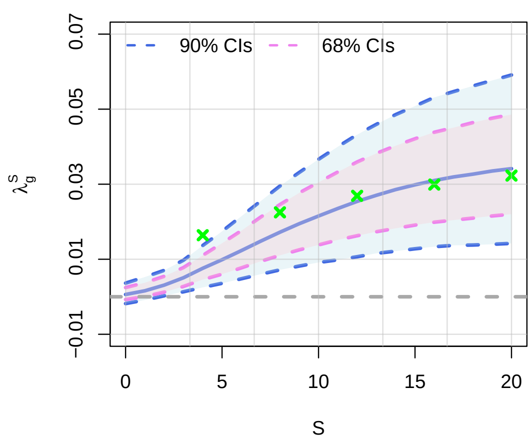
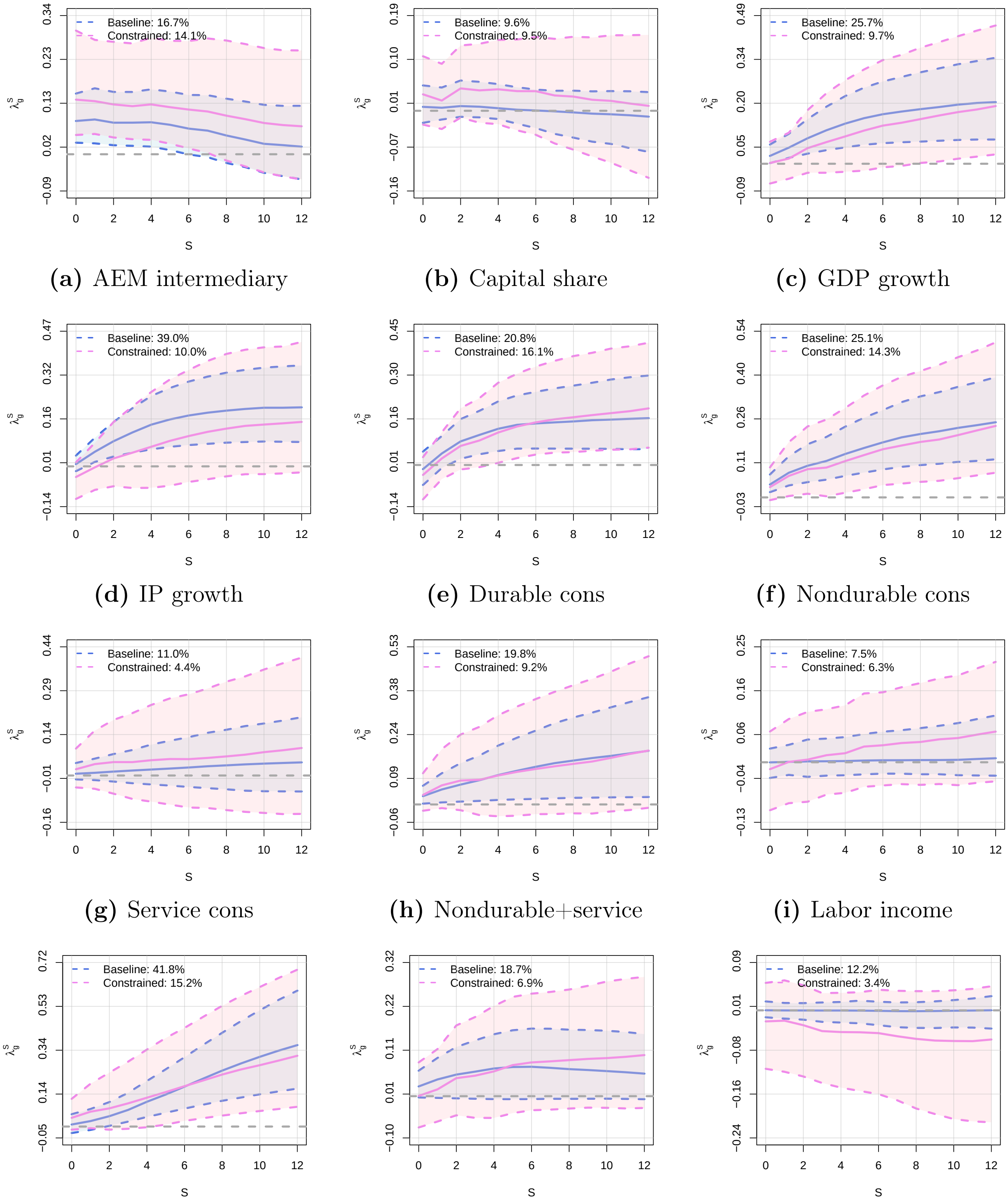



**Figure IA.2:** Term structure of unconditional dividend risk premia

Term structure of dividend risk premia (not standardized, unlike Table 1). Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns using 20-quarter lags in $g_t$'s equations. Pink and blue shaded areas show 68% and 90% Bayesian credible intervals, respectively. Green crosses: risk premia estimates from Bansal et al. (2021), Table 4.



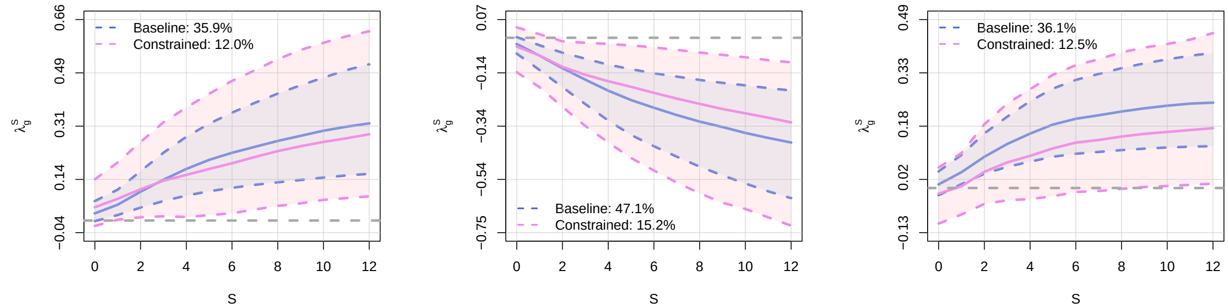

**(a)** AEM intermediary **(b)** Capital share **(c)** GDP growth

**(d)** IP growth **(e)** Durable cons **(f)** Nondurable cons

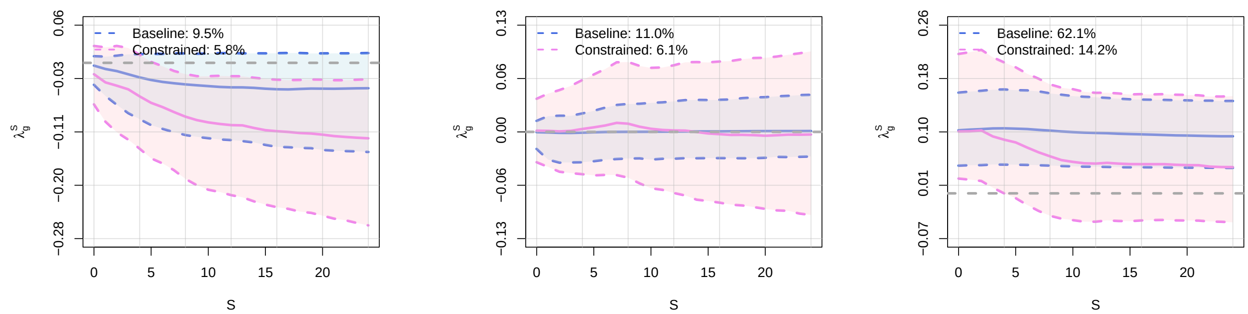
**(g)** Service cons **(h)** Nondurable+service **(i)** Labor income

**(j)** Dividend growth **(k)** TFP growth **(l)** TFP growth (utilization)

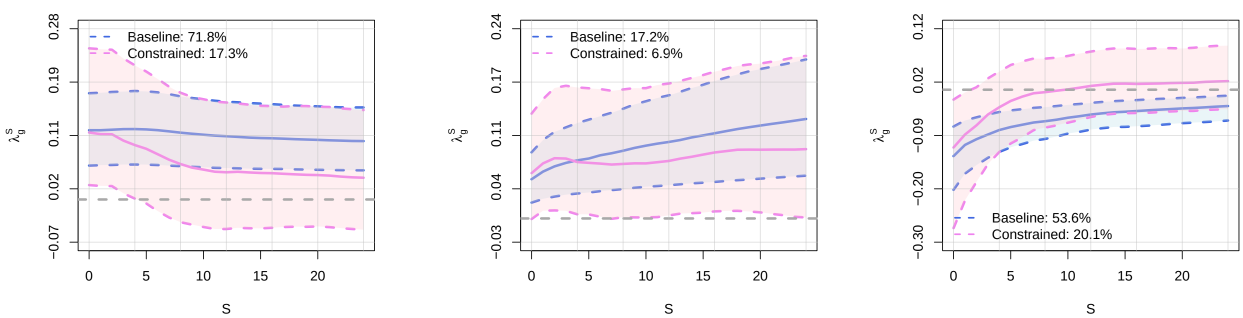
**Figure IA.3:** Term structure of unconditional risk premia in five-factor models: Baseline vs. restricted models




**(m)** Hours worked growth **(n)** Δ Unemployment rate **(o)** Investment growth



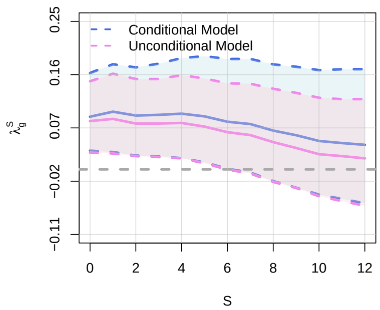

**(p)** Oil price change **(q)** TED spread **(r)** $HKM_{ntr}$



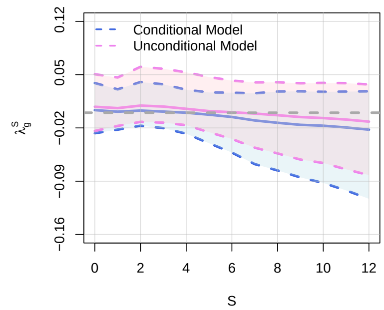

**(s)** $HKM_{tr}$ **(t)** PS liquidity **(u)** $\Delta \log(\text{VIX})$

**Figure IA.3:** Term structure of unconditional risk premia in five-factor models: Baseline vs. restricted models

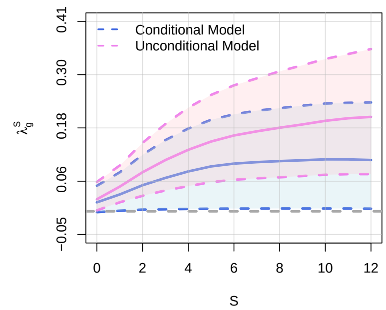
Term structure of unconditional risk premia: baseline vs. restricted models. Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns. For quarterly (monthly) factors, we use 12-quarter (24-month) lags in $g_t$'s equations. Blue lines: baseline model ($g_t$ projected onto latent factors). Purple lines: restricted model ($g_t$ projected onto SDF). Shaded areas: 90% credible intervals. Time-series $R^2$s shown for both models. Data sources in Appendix IA.4.



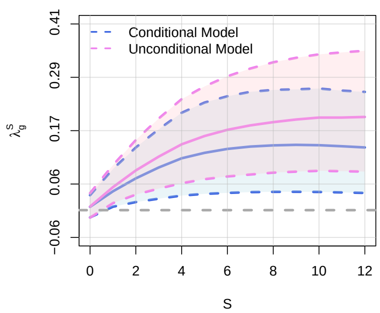

**(a)** AEM intermediary



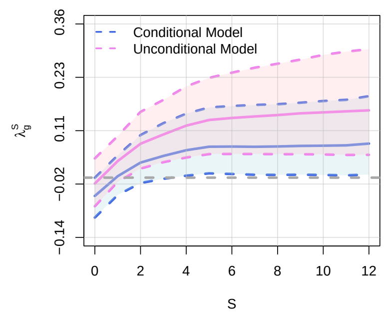

**(b)** Capital share



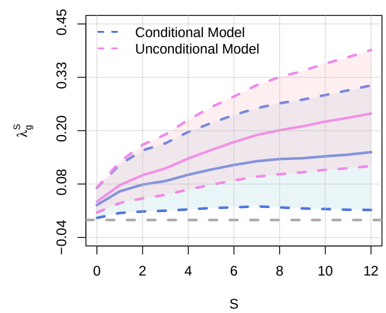

**(c)** GDP growth



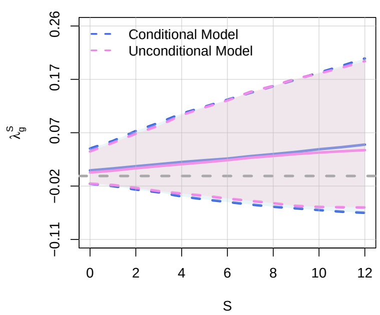

**(d)** IP growth



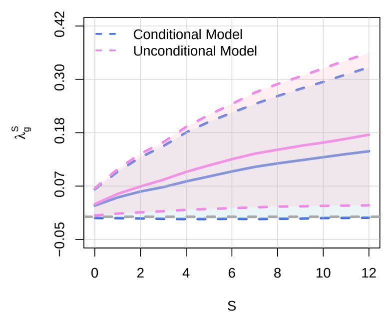

**(e)** Durable cons



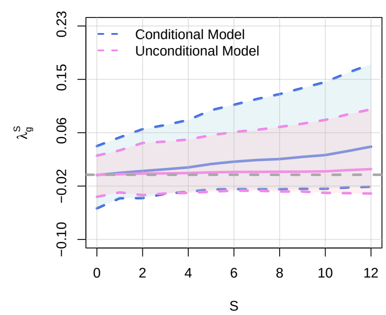

**(f)** Nondurable cons



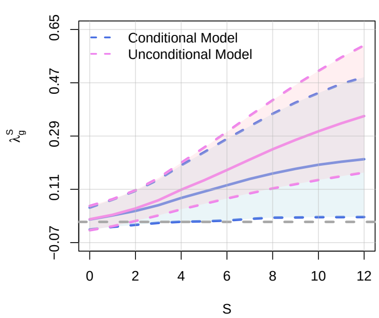

**(g)** Service cons



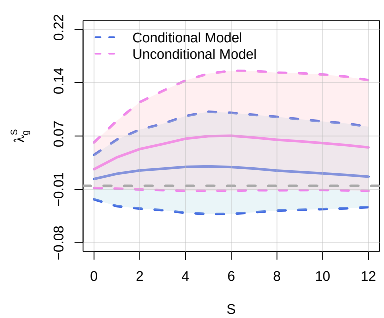

**(h)** Nondurable+service



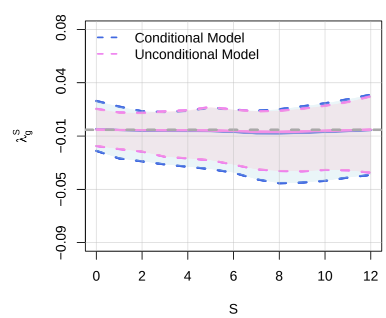

**(i)** Labor income




**(j)** Dividend growth

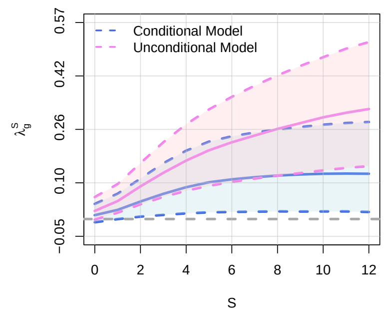



**(k)** TFP growth

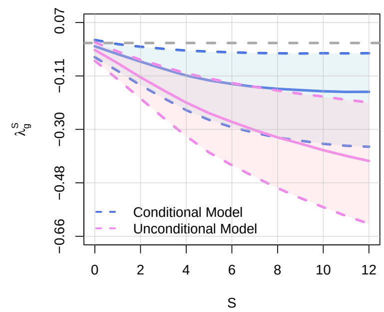



**(l)** TFP growth (utilization)

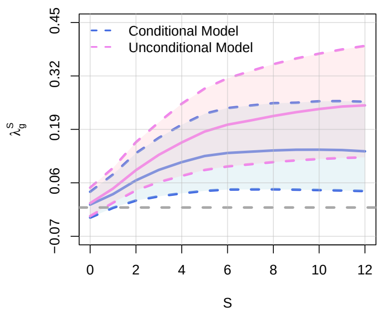
**Figure IA.4:** Term structure of unconditional risk premia in time-varying models. External predictors: PE ratio of S&P 500, Term spread, default spread, and value spread



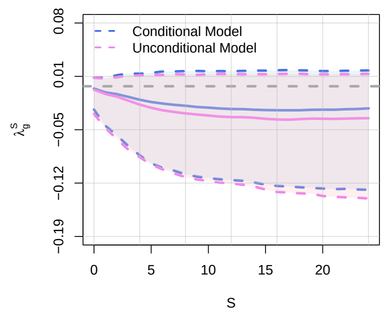

**(m)** Hours worked growth



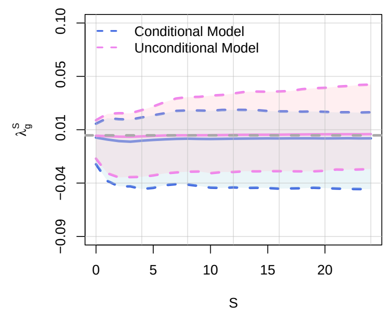

**(n)** $\Delta$ Unemployment rate



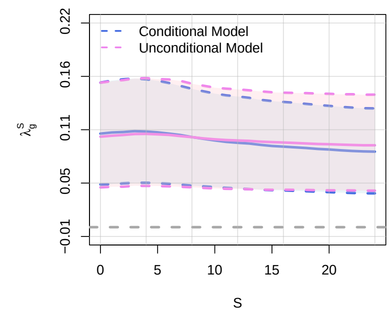

**(o)** Investment growth



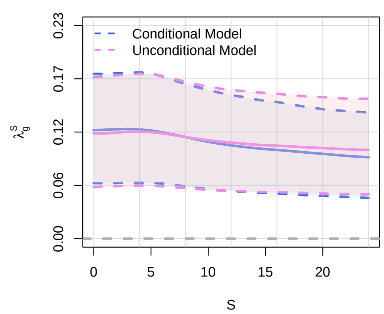

**(p)** Oil price change



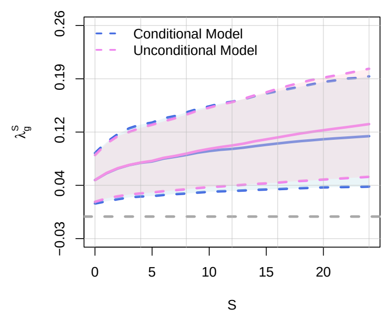

**(q)** TED spread



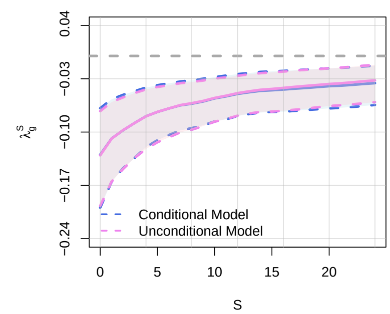

**(r)** $HKM_{ntr}$




**(s)** $HKM_{tr}$




**(t)** PS liquidity




**(u)** $\Delta\log$(VIX)

**Figure IA.4:** Term structure of unconditional risk premia in time-varying models. External predictors: PE ratio of S&P 500, Term spread, default spread, and value spread

Term structure of unconditional risk premia using Propositions 1 and A1. Base assets: 275 Fama-French characteristic-sorted portfolios. Five-factor model for returns. For quarterly (monthly) factors, we use 12-quarter (24-month) lags in $g_t$'s equations. Blue lines and shaded areas: conditional model estimates with 90% credible intervals (Section 1.2), where latent factors follow a VAR(1) with both factors and four external predictors (PE ratio of S&P 500, term spread, default spread, value spread) driven by lagged predictors. Purple lines and shaded areas: unconditional model estimates (Section 1.1). Data sources in Appendix IA.4.