Macroeconomic Determinants of Stock Market Volatility and Volatility Risk-Premiums

Valentina Corradi[a], Walter Distaso[b], Antonio Mele[c,*,†]

[a] *University of Warwick;* [b] *Imperial College Business School;* [c] *London School of Economics*

August 8, 2011.

Abstract

How does stock market volatility relate to the business cycle? We develop, and estimate, a no-arbitrage model to study the cyclical properties of stock volatility and the risk-premiums the market requires to bear the risk of fluctuations in this volatility. The level and fluctuations of stock market volatility can be largely explained by business cycle factors, although some unobserved factor contributes to nearly 20% to the overall variation in volatility. At the same time, this unobservable factor cannot explain the ups and downs volatility experiences over time—the "volatility of volatility." Instead, the volatility of volatility relates to the business cycle. Finally, volatility risk-premiums are strongly countercyclical, even more so than stock volatility, and are partially responsible for the large swings in the VIX index occurred during the 2007-2009 subprime crisis, which our model does capture in out-of-sample experiments.

Keywords: Aggregate stock market volatility; volatility risk-premiums; volatility of volatility; business cycle; no-arbitrage restrictions; simulation-based inference

JEL classification: E37, E44, G13, G17, C15, C32

*We wish to thank Urban Jermann (the co-editor), two anonymous referees, and Yacine Aït-Sahalia, Alessandro Beber, Bernard Dumas, Marcelo Fernandes, Daniel Gutknecht, Christian Julliard, Nour Meddahi, Mark Salmon, George Tauchen, Viktor Todorov, and seminar participants at EDHEC-Nice, European Central Bank, HEC Lausanne, Hong Kong Monetary Authority, IE Business School (Madrid), the Universities of Aarhus (CREATES), Bicocca (Milan), Goethe (Frankfurt), Louvain, Lugano (Finance) and Warwick, the 2006 London-Oxford Financial Econometrics Study Group, the 2006 conference on Realized Volatility (Montréal), the 2008 UCL-CEMMAP conference, the 2008 Imperial College Financial Econometrics Conference, the 2008 Adam Smith Asset Pricing Conference at London Business School, the 2008 LSE-FMG conference on "Integrating historical data and market expectations in financial econometrics," the 2008 North American Summer Meeting of the Econometric Society (Carnegie), and the 2008 SITE Summer Workshop at Stanford University for valuable comments and suggestions. Valentina Corradi and Walter Distaso acknowledge support from the British ESRC under the grant RES-062-23-0311. Antonio Mele thanks the British EPSRC for support via grant EP/C522958/1. The usual disclaimer applies.

†Corresponding author.

1. Introduction

Understanding the origins of stock market volatility has long been a topic of considerable interest to both policy makers and market practitioners. Policy makers are interested in the main determinants of volatility and in its spillover effects on real activity. Market practitioners are interested in the effects volatility exerts on the pricing and hedging of plain vanilla options and more exotic derivatives. In both cases, forecasting stock market volatility constitutes a formidable challenge but also a fundamental instrument to manage the risks faced by these institutions.

Many available models use latent factors to explain the dynamics of stock market volatility. For example, in the celebrated Heston's (1993) model, stock volatility is exogenously driven by some unobservable factor correlated with the asset returns. Yet such an unobservable factor does not bear an economic interpretation. Moreover, the model implies, by assumption, that volatility cannot be forecasted by macroeconomic factors such as industrial production or inflation. This circumstance is counterfactual. Indeed, there is strong evidence that stock market volatility has a very pronounced business cycle pattern, being higher during recessions than during expansions; see, e.g., Schwert (1989a,b), Hamilton and Lin (1996), or Brandt and Kang (2004).

In this paper, we develop a no-arbitrage model where stock market volatility is explicitly related to a number of macroeconomic and unobservable factors. The distinctive feature of this model is that stock volatility is linked to these factors by no-arbitrage restrictions. The model is also analytically convenient: under fairly standard conditions on the dynamics of the factors and risk-aversion corrections, our model is solved in closed-form, and is amenable to empirical work.

We use the model to quantitatively assess how market volatility and volatility-related risk-premiums change in response to business cycle conditions. Our model fully captures the procyclical nature of aggregate returns and the countercyclical behavior of stock volatility that we have been seeing in the data for a long time. It makes a fundamental prediction: macroeconomic factors can explain nearly 75% of the variation in the overall stock volatility. At the same time, our model, rigorously estimated through simulation-based inference methods, shows that the presence of some unobservable and persistent factor is needed to sustain the level of stock volatility that matches its empirical counterpart. Moreover, our model reveals that macroeconomic factors substantially help explain the variability of stock volatility around its level—the volatility of volatility. That such a "vol-vol" might be related to the business cycle is indeed a plausible hypothesis, although clearly, the ups and downs stock volatility experiences over the business cycle are a prediction of the model in line with the data, not a restriction imposed while estimating the model. Such a new property we uncover, and model, brings practical implications. For example, business cycle forecasters might learn that not only does stock market volatility have predicting power, as discussed below; "vol-vol" is also a potential predictor of the business cycle.

Our second set of results relates to volatility-related risk-premiums. The volatility risk-premium is the difference between the expectation of future market volatility under the risk-neutral and the true probability. It quantifies how a representative agent is willing to pay to be ensured against the event that volatility will raise beyond his own expectations. Thus, it is a very intuitive and general measure of risk-aversion. We find that this volatility risk-premium is strongly countercyclical, even more so than stock volatility. Precisely, volatility risk-premiums are typically not very volatile, although in bad times, they may increase to extremely high levels, and quite quickly. We undertake a stress test of the model over a particularly uncertain period, which includes the 2007-2009 subprime turmoil. Ours is a stress test, as (i) we estimate the model using post-war data up to 2006, and (ii) feed the previously estimated model with macroeconomic data related to the subprime crisis. We compare the model's predictions for the crisis with the actual behavior of both stock volatility and the new VIX index, maintained by the Chicago Board Options Exchange (CBOE), which is, theoretically, the risk-adjusted expectation of future volatility within one month. The model tracks the dramatic movements in this index, and predicts that countercyclical volatility risk-premiums are largely responsible for the large swings in the VIX occurred during the crisis. In fact, we show that over this crisis, as well as in previous recessions, movements in the VIX index are determined by changes in such countercyclical risk-premiums, not by changes in the expected volatility.

Related literature

Stock volatility and volatility risk-premiums The cyclical properties of aggregate stock market volatility have been the focus of recent empirical research, although early work relating stock volatility to macroeconomic variables dates back to King, Sentana and Wadhwani (1994), who rely on a no-arbitrage model. In a comprehensive international study, Engle and Rangel (2008) find that high frequency aggregate stock volatility has both a short-run and long-run component, and suggest that the long-run component is related to the business cycle. Adrian and Rosenberg (2008) show that the short- and long- run components of aggregate volatility are both priced, cross-sectionally. They also relate the long-run component of aggregate volatility to the business cycle. Finally, Campbell, Lettau, Malkiel and Xu (2001), Bloom (2009), Bloom, Floetotto and Jaimovich (2009) and Fornari and Mele (2010) show that capital market uncertainty helps explain future fluctuations in real economic activity. Our focus on the volatility risk-premiums relates, instead, to the seminal work of Dumas (1995), Bakshi and Madan (2000), Britten-Jones and Neuberger (2000), and Carr and Madan (2001), which has more recently stimulated an increasing interest in the dynamics and determinants of the volatility risk-premium (see, for example, Bakshi and Madan (2006) and Carr and Wu (2009)). Notably, in seminal work, Bollerslev and Zhou (2006) and Bollerslev, Gibson and Zhou (2011) unveil, empirically, a strong relation between this volatility risk-premium and a number of macroeconomic factors.

Our contribution hinges upon, and expands, over this growing literature, in that we formulate and estimate a fully-specified no-arbitrage model relating the dynamics of stock volatility and volatility risk-premiums to business cycle, and additional unobservable, factors. With the exception of King, Sentana and Wadhwani (1994) and Adrian and Rosenberg (2008), who still have a focus different from ours, the predicting relations in the previous papers, while certainly useful, are still part of reduced-form statistical models. In our out-of-sample experiments of the subprime crisis, we shall show that our no-arbitrage framework is considerably richer than that based on predictive linear regressions. We show, for example, that compared to our model's predictions about stock volatility and the VIX index, predictions from linear regressions are substantially flat over the subprime crisis.

The only antecedent to our paper is Bollerslev, Tauchen and Zhou (2009), who develop a consumption-based rationale for volatility risk-premiums, although then, the authors use this rationale only as a guidance to the estimation of reduced-form predictability regressions conditioned on the volatility risk-premium. In recent independent work discussed below, Drechsler and Yaron (2011) investigate the properties of the volatility risk-premium, implied by a calibrated consumption-based model with long-run risks. The authors, however, are not concerned with the cross-equation restrictions relating the volatility risk-premium to state variables driving low frequency stock market fluctuations which, instead, constitute the central topic of our paper.

No-arbitrage regressions In recent years, there has been a significant surge of interest in consumption-based explanations of aggregate stock market volatility (see, for example, Campbell and Cochrane (1999), Bansal and Yaron (2004), Tauchen (2005), Mele (2007), or the two surveys in Campbell (2003) and Mehra and Prescott (2003)). These explanations are important because they highlight the main economic mechanisms through which markets and preferences affect equilibrium asset prices and, hence, stock volatility. In our framework, cross-equations restrictions arise through the weaker requirement of absence of arbitrage opportunities. In this respect, our approach is similar in spirit to the "no-arbitrage" vector autoregressions introduced in the term-structure literature by Ang and Piazzesi (2003) and Ang, Piazzesi and Wei (2006). Similarly as in those papers, we specify an analytically convenient pricing kernel affected by some macroeconomic factors, although we do not directly relate these to, say, markets, preferences or technology.

Our model works quite simply. We exogenously specify the joint dynamics of a number of macroeconomic and unobservable factors. We assume that the asset payoffs and the risk-premiums required by agents to be compensated for the fluctuations of the factors, are essentially affine functions of these factors, along the lines of Duffee (2002). We show that the resulting no-arbitrage stock price is affine in the factors. Our model does not allow for jumps or other market micro-structure effects, as our main focus is to model low frequency movements in the aggregate stock volatility and volatility risk-premiums, through the use of macroeconomic and unobservable factors. Our estimation results, obtained through data sampled at monthly frequency, are unlikely to be affected by measurement noise or jumps, say. In related work, Carr and Wu (2009), Todorov (2010), Drechsler and Yaron (2011), and Todorov and Tauchen (2011) do allow for the presence of jumps, although they do not analyze the relations between macroeconomic variables and aggregate volatility or volatility risk-premiums, which we do here.

Estimation strategy, and plan of the paper

In standard stochastic volatility models such as that in Heston (1993), volatility is driven by factors, which are not necessarily the same as those affecting the stock price—volatility is exogenous in these models. In our no-arbitrage model, volatility is endogenous, relating to a number of risks affecting (i) macroeconomic developments, (ii) unobserved factors and (iii) the very same asset returns—these risks affect both asset returns and volatility. To identify the premium required to bear the risk of volatility, we exploit derivative data, related to the new VIX index.

We implement a three-step estimation procedure that relies on simulation-based inference methods. In the first step, we estimate the parameters underlying the macroeconomic factors. In the second step, we use data on a broad stock market index, and the macroeconomic factors, and estimate reduced-form parameters linking the stock market index to the macroeconomic factors and the third unobservable factor, as well as the parameters underlying the dynamics of the unobservable factor. In the third step, we use data on the new VIX index, and the macroeconomic factors, and estimate the risk-premiums parameters. We implement these steps by matching model-based moments and impulse response functions to their empirical counterparts, relating to macroeconomic factors, realized returns, realized volatility and the VIX index. We develop, and utilize, a theory to consistently estimate the standard errors through block-bootstrap methods.

The remainder of the paper is organized as follows. In Section 2 we develop a no-arbitrage model for the stock price, stock volatility and volatility-related risk-premiums. Section 3 illustrates the estimation strategy. Section 4 presents our empirical results. Section 5 concludes, and the Supplemental material contains an appendix with technical details omitted from the main text.

2. The model

We develop a model where aggregate stock returns and volatility are tied up to macroeconomic developments and one unobservable factor. It is a three-factor model solved in closed form, a special case of a general multifactor model in Appendix A of the Supplemental material.

2.1. The macroeconomic environment

We consider a model with one unobservable factor, and two additional factors affecting the development of two aggregate macroeconomic variables, inflation and industrial production growth, and the stock market. Let $\boldsymbol{y}(t) = (y_1(t)\ \ y_2(t)\ \ y_3(t))$ be a vector-valued process, where $y_1(t)$ and $y_2(t)$ denote two observable factors, defined as $\ln(\mathrm{CPI}_t/\mathrm{CPI}_{t-12}) = \ln y_1(t)$ and $\ln(\mathrm{IP}_t/\mathrm{IP}_{t-12}) = \ln y_2(t)$, where CPI_t and IP_t are the consumer price index and industrial production as of month t, as further explained in Section 4.1. In Section 4.1, we also discuss the role these two macroeconomic factors have played in asset pricing. We also assume that a third, and unobservable, factor, $y_3(t)$, affects the stock price, but not the two macroeconomic aggregates, CPI_t and IP_t. Finally, we assume the two macroeconomic factors do not affect the unobservable factor y_3, although we allow for simultaneous feedback effects between inflation and industrial production growth, as explained below. The factors y_j are solution to,

$$\mathrm{d}y_j(t) = \left[\kappa_j\left(\mu_j - y_j(t)\right) + \bar{\kappa}_j\left(\bar{\mu}_j - \bar{y}_j(t)\right)\right]\mathrm{d}t + \sqrt{\alpha_j + \beta_j y_j(t)}\,\mathrm{d}W_j(t), \quad j = 1, 2, 3, \tag{1}$$

where $W_j(t)$ are standard Brownian motions, $\bar{\mu}_1 \equiv \mu_2$, $\bar{y}_1(t) \equiv y_2(t)$, $\bar{\mu}_2 \equiv \mu_1$, $\bar{y}_2(t) = y_1(t)$, $\bar{\kappa}_3 \equiv \bar{\mu}_3 \equiv \bar{y}_3(t) \equiv 0$ and, finally, Greek letters denote constant parameters. The two parameters, κ_1 and κ_2, are the speed of adjustment of inflation and industrial production growth towards their long run means, μ_1 and μ_2, and $\bar{\kappa}_1$ and $\bar{\kappa}_2$ are the feedback parameters. Appendix A of the Supplemental material reviews conditions guaranteeing Eq. (1) is well-defined, which we use as constraints whilst estimating the model.

We assume that asset prices, (i) respond to movements in the factors affecting macroeconomic conditions, and (ii) reflect a long-run trend in the asset payoffs. Precisely, we model the instantaneous dividends paid off by the asset at time t, $\mathrm{Div}(t)$ say, as the product of a stochastic trend, times a stationary component, as follows:

$$\text{Div}\,(t) = G\,(t)\,\delta\,(\boldsymbol{y}\,(t))\,, \tag{2}$$

where $\delta\,(\boldsymbol{y})$ satisfies, for four constants δ_0 and $(\delta_j)_{j=1}^{3}$,

$$\delta\,(\boldsymbol{y}) = \delta_0 + \delta_1 y_1 + \delta_2 y_2 + \delta_3 y_3, \tag{3}$$

and $G\,(t)$ is a geometric Brownian motion with drift g and volatility σ_G,

$$\frac{dG\,(t)}{G\,(t)} = gdt + \sigma_G dW_G\,(t)\,, \quad G\,(0) \equiv 1, \tag{4}$$

and $W_G\,(t)$ is a Brownian motion uncorrelated with the Brownian motions in Eq. (1). The rationale behind the assumption in Eq. (2) is to disentangle secular, yet stochastic, dividend growth, captured by $G\,(t)$, from short-run fluctuations of the dividend process, arising from business cycles, and captured by $\delta\,(\boldsymbol{y}\,(t))$. This assumption implies the asset price displays a similar property, being driven by a secular, growth component, and an additional, short-run component related to macroeconomic developments, as we now explain.

2.2. No-arbitrage

We model the pricing kernel, or the Arrow-Debreu price density, in the economy. Let $\mathbb{F}\,(T)$ be the sigma-algebra generated by the Brownian motion $[\boldsymbol{W}\,(t)^{\top} \;\; W_G\,(t)]^{\top}$, $t \leq T$, where $\boldsymbol{W}\,(t) = (y_1\,(t) \;\; y_2\,(t) \;\; y_3\,(t))$, and let P the associated physical probability. The Radon-Nikodym derivative of the risk-neutral probability Q with respect to P on $\mathbb{F}\,(T)$ is,

$$\xi(T) \equiv \frac{\mathrm{d}Q}{\mathrm{d}P} = \exp\left(-\int_0^T \boldsymbol{\Lambda}\,(t)^{\top}\,\mathrm{d}\boldsymbol{W}\,(t) - \frac{1}{2}\int_0^T \|\boldsymbol{\Lambda}\,(t)\|^2\,\mathrm{d}t\right)\cdot\exp\left(-\lambda_G W_G\,(T) - \frac{1}{2}\lambda_G^2 T\right), \tag{5}$$

for some risk-premium process $\boldsymbol{\Lambda}\,(t)$ and constant λ_G. The interpretation of $\boldsymbol{\Lambda}\,(t)$ is that of a risk-premium required to compensate for the fluctuations of the factors $\boldsymbol{y}\,(t)$. The constant λ_G is, instead, the unit-risk premium for the stochastic fluctuations of secular growth, $G\,(t)$. While we model $\boldsymbol{\Lambda}\,(t)$ to be time-varying, we assume λ_G to be constant for analytical convenience.

We assume the risk-premium process satisfies an "essentially affine" specification, viz

$$\boldsymbol{\Lambda}\,(\boldsymbol{y}\,(t)) \equiv \boldsymbol{\Lambda}\,(t) = \boldsymbol{V}\,(\boldsymbol{y}\,(t))\,\boldsymbol{\lambda}_1 + \boldsymbol{V}^{-}\,(\boldsymbol{y}\,(t))\,\boldsymbol{\lambda}_2\boldsymbol{y}\,(t)\,, \tag{6}$$

where $\boldsymbol{\lambda}_1 = \left(\lambda_{1(1)}\ \lambda_{1(3)}\ \lambda_{1(3)}\right)$ is a parameter vector, $\boldsymbol{\lambda}_2$ is a diagonal matrix of parameters with diagonal elements equal to $\lambda_{2(j)}$, $j = 1, 2, 3$, $\boldsymbol{V}(\boldsymbol{y})$ is a diagonal matrix with $\sqrt{\alpha_j + \beta_j y_j}$ on its diagonal, and $\boldsymbol{V}^{-}(\boldsymbol{y}) : \boldsymbol{V}^{-}(\boldsymbol{y})\,\boldsymbol{V}(\boldsymbol{y}) = \boldsymbol{I}_{3\times 3}$, for all $\boldsymbol{y}$, which it does under regularity conditions spelled out in Appendix A of the Supplemental material.

The functional form for $\boldsymbol{\Lambda}$ echoes that suggested by Duffee (2002) in the term-structure literature. If $\boldsymbol{\lambda}_2 = \boldsymbol{0}_{3\times 3}$, $\boldsymbol{\Lambda}$ collapses to the "completely affine" specification introduced by Duffie and Kan (1996), where the risk-premiums in $\boldsymbol{\Lambda}$ are tied up to the volatility of the fundamentals, $\boldsymbol{V}(\boldsymbol{y})$. While it is reasonable to assume that risk-premiums link to the *volatility* of fundamentals, the specification in Eq. (6) also allows risk-premiums to relate to the *level* of the fundamentals, through the additional term $\boldsymbol{\lambda}_2\boldsymbol{y}$. Including this term is, indeed, critical to our empirical results. Consider the *total* risk-premiums process, defined as,

$$\boldsymbol{\pi}(\boldsymbol{y}) = \begin{pmatrix} \pi_1(y_1) \\ \pi_2(y_2) \\ \pi_3(y_3) \end{pmatrix} \equiv \boldsymbol{V}(\boldsymbol{y})\,\boldsymbol{\Lambda}(\boldsymbol{y}) = \begin{pmatrix} \alpha_1\lambda_{1(1)} + \left(\beta_1\lambda_{1(1)} + \lambda_{2(1)}\right) y_1 \\ \alpha_2\lambda_{1(2)} + \left(\beta_2\lambda_{1(2)} + \lambda_{2(2)}\right) y_2 \\ \alpha_3\lambda_{1(3)} + \left(\beta_3\lambda_{1(3)} + \lambda_{2(3)}\right) y_3 \end{pmatrix}. \tag{7}$$

Each component of $\boldsymbol{\pi}(\boldsymbol{y})$, $\pi_j(y_j)$, depends on factor y_j due to the volatility of this factor (i.e. through β_j) and, also, due to the additional parameter $\lambda_{2(j)}$. Without $\lambda_{2(j)}$, we could not model the level of the risk-premiums separately from their sensitivities to changes in y_j—a sensible issue we have experienced whilst estimating our model. Consider, for example, the total risk premium for growth, $\pi_2(y_2)$. The coefficient $\lambda_{1(2)}$ affects both the intercept and the slope of π_2. The inclusion of $\lambda_{2(2)}$ allows to achieve flexibility in modeling the level of $\pi_2(y_2)$ and its sensitivity with respect to changes in y_2.

Finally, we assume that the safe asset is elastically supplied such that the short-term rate r (say) is constant. Whilst real rates are not as volatile as stock returns in the data, many existing models might likely predict rates to be too volatile. For example, models with habit formation predict the short-term rate is a function of the state, primarily due to intertemporal substitution effects. Campbell and Cochrane (1999) mitigate this issue with a well-known trick—they impose that intertemporal substitution effects are exactly offset by precautionary savings, thereby making the short-term rate constant. Additional models that cope with this challenge include those relying on non-expected utility, as in Bansal and Yaron (2004), or those with heterogeneous agents, as in Guvenen (2009), to cite a few. In this paper, we impose r to be constant for the purpose of keeping stock volatility tractable, as this facilitates the actual estimation of the model. How important is this assumption, quantitatively? Mele (2007) finds that in realistically calibrated models of habit formation, large countercyclical swings of stock volatility mainly arise due to risk-premiums effects, rather than interest rate volatility. It is an open question, however, whether such a result would still hold in the economy we consider in the current paper.[1]

We are ready to determine the no-arbitrage stock price. As it turns out, the previous assumption on the pricing kernel and the assumption that $\delta(\cdot)$ in Eq. (3) is affine in $\boldsymbol{y}$ implies that the stock price is also affine in $\boldsymbol{y}$. Precisely, we have:

$$S(G, \boldsymbol{y}) = G \cdot \left(s_0 + \sum\nolimits_{j=1}^{3} s_j y_j \right), \tag{8}$$

where

$$s_0 = \frac{1}{r - g + \sigma_G \lambda_G} \left[\delta_0 + \sum\nolimits_{j=1}^{3} s_j \left(\kappa_j \mu_j + \bar{\kappa}_j \bar{\mu}_j - \alpha_j \lambda_{1(j)} \right) \right], \tag{9}$$

$$s_j = \frac{\delta_j \left(r - g + \sigma_G \lambda_G + \kappa_i + \lambda_{1(i)} \beta_i + \lambda_{2(i)} \right) - \delta_i \bar{\kappa}_i}{\prod_{h=1}^{2} \left(r - g + \sigma_G \lambda_G + \kappa_h + \lambda_{1(h)} \beta_h + \lambda_{2(h)} \right) - \bar{\kappa}_1 \bar{\kappa}_2}, \quad \text{for } j, i \in \{1, 2\} \text{ and } i \neq j, \tag{10}$$

$$s_3 = \frac{\delta_3}{r - g + \sigma_G \lambda_G + \kappa_3 + \lambda_{1(3)} \beta_3 + \lambda_{2(3)}}. \tag{11}$$

In the standard stochastic volatility literature, the asset price and, hence, its volatility, is taken as given, and volatility and volatility risk-premiums are modeled separately, as for example in the celebrated Heston's (1993) model, which many empirical studies take as a benchmark (e.g., Chernov and Ghysels (2000), Corradi and Distaso (2006), Garcia, Lewis, Pastorello and Renault (2011)). Moreover, a recent focus in this literature is to relate volatility risk-premiums the to business cycle (e.g., Bollerslev, Gibson and Zhou (2011)). Yet, while the empirical results in these papers are ground breaking, the Heston's model is not meant to capture, theoretically, the interplay between stochastic volatility, volatility risk-premiums and the business cycle.

[1] Our model has, however, implications for the nominal rate, which is $r - \ln \mathbb{E}\left(\frac{\mathrm{CPI}_t}{\mathrm{CPI}_{t+12}} \middle| \mathbb{F}_t \right)$ (for one year, say), where $\mathbb{E}$ is the expectation under Q. Evaluating this expression in steady state, through the estimates we obtain in Section 4, and assuming $r = 1\%$, yields 4.7%. In the data, the nominal rate for one year is, instead, 5.4%.

Our model works differently, as it places restrictions on the asset price process directly, through our assumptions on the fundamentals of the economy, and absence of arbitrage. For our model, it is the asset price that determines, endogenously, volatility, which by Eq. (1) and Eq. (8) is:

$$\sigma\left(\boldsymbol{y}\left(t\right)\right) \equiv \sigma\left(t\right) = \sqrt{\sigma_G^2 + \frac{\sum_{j=1}^{3} s_j^2\left(\alpha_j + \beta_j y_j\left(t\right)\right)}{\left(s_0 + \sum_{j=1}^{3} s_j y_j\left(t\right)\right)^2}}. \tag{12}$$

Note that the model predicts that stock volatility embeds information about risk-corrections that agents require to invest in the stock market. We shall make use of this observation in the empirical part of the paper. We now describe which measure of stock volatility we use to proceed with such a critical step of our analysis.

2.3. Arrow-Debreu adjusted volatility

In September 2003, the CBOE changed its volatility index VIX, to reflect recent advances in the option pricing literature. Given an asset price process $S\left(t\right)$ that is continuous in time (as that predicted by our model, in Eq. (8)), and all available information $\mathbb{F}\left(t\right)$ at time t, consider the economic value of the future integrated variance on a given interval $[t, t_0]$, IV_{t,t_0}, say, which is the sum of the future variances, weighted with the Arrow-Debreu state prices:

$$\mathbb{E}\left[IV_{t,t_0}\middle|\,\mathbb{F}\left(t\right)\right] \equiv \int_t^{t_0} \mathbb{E}\left[\left(\frac{\mathrm{d}}{\mathrm{d}\tau}\mathrm{var}\left[\ln S\left(\tau\right)\middle|\,\mathbb{F}\left(u\right)\right]\Big|_{\tau=u}\right)\middle|\,\mathbb{F}\left(t\right)\right]\mathrm{d}u, \tag{13}$$

where $\mathbb{E}$ is the expectation under Q. The new VIX index relies on the work of Dumas (1995), Bakshi and Madan (2000), Britten-Jones and Neuberger (2000), and Carr and Madan (2001), who showed that the risk-neutral expectation of the future integrated variance is a functional of put and call options written on the asset:

$$\mathbb{E}\left[IV_{t,t_0}\middle|\,\mathbb{F}\left(t\right)\right] = 2e^{r(t_0-t)}\left[\int_0^{F(t)}\frac{P_t\left(t_0,K\right)}{K^2}\mathrm{d}K + \int_{F(t)}^{\infty}\frac{C_t\left(t_0,K\right)}{K^2}\mathrm{d}K\right] \equiv \left(t_0 - t\right)\cdot \mathrm{VIX}_t^2, \tag{14}$$

where $F\left(t\right) = e^{r(t_0-t)}S\left(t\right)$ is the forward price, $C_t\left(t_0,K\right)$ and $P_t\left(t_0,K\right)$ are the prices as of time t of call and put options expiring at t_0 and struck at K, and VIX_t is the new VIX index. In contrast, our model, which relies on the Arrow-Debreu state prices in Eq. (5), predicts that the risk-neutral expectation of the integrated variance is:

$$\mathbb{E}\left[IV_{t,t_0}\middle|\,\boldsymbol{y}\left(t\right)=\boldsymbol{y}\right] = \int_t^{t_0}\mathbb{E}\left[\sigma^2\left(\boldsymbol{y}\left(u\right)\right)\middle|\,\boldsymbol{y}\left(t\right)=\boldsymbol{y}\right]\mathrm{d}u \equiv \left(t_0-t\right)\cdot \mathrm{VIX}^2\left(\boldsymbol{y}\right), \tag{15}$$

where $\sigma^2(\boldsymbol{y}(t))$ is given in Eq. (12). We shall estimate the risk-premium parameters in Eq. (7) so as to match the VIX index predicted by the model, $\text{VIX}(\boldsymbol{y}(t))$ in Eq. (15), to its empirical counterpart, VIX_t in Eq. (14). Finally, our model makes predictions about how the volatility risk-premium, $\text{VRP}(\boldsymbol{y}(t))$ say, changes with the factors $\boldsymbol{y}(t)$ in Eq. (1)

$$\text{VRP}(\boldsymbol{y}(t)) \equiv \sqrt{\frac{1}{t_0 - t}} \left(\sqrt{\mathbb{E}\left[IV_{t,t_0} \middle| \boldsymbol{y}(t) = \boldsymbol{y} \right]} - \sqrt{E\left[IV_{t,t_0} \middle| \boldsymbol{y}(t) = \boldsymbol{y} \right]} \right), \tag{16}$$

where E denotes the expectation taken under P.

3. Statistical inference

We rely on a three-step procedure. In the first step, we estimate the parameters of the process underlying the dynamics of the two macroeconomic factors, $\boldsymbol{\phi}^\top = \left(\kappa_j, \mu_j, \alpha_j, \beta_j, \bar{\kappa}_j, j = 1, 2\right)$. In the second step, we estimate the parameters in Eq. (4), $\boldsymbol{\theta}_G^\top = (g, \sigma_G)$, the reduced-form parameters that link the asset price to the three factors in Eq. (8), and the parameters of the process for the unobserved factor, $\boldsymbol{\theta}^\top = \left(\kappa_3, \mu_3, \alpha_3, \beta_3, (s_j)_{j=0}^3\right)$, while imposing the identifiability condition that $\mu_3 = 1$, as explained below. In the third step, we estimate the risk-premiums parameters $\boldsymbol{\lambda}^\top = \left(\lambda_{1(1)}, \lambda_{2(1)}, \lambda_{1(2)}, \lambda_{2(2)}, \lambda_{1(3)}, \lambda_{2(3)}\right)$, relying on a simulation-based approximation of the model-implied VIX, which we match to the VIX index. At each of these steps, we do not have a closed form expression of either the likelihood function or selected sets of moment conditions. For this reason, we need to rely on a simulation-based approach. Our estimation strategy relies on an hybrid of Indirect Inference (Gouriéroux, Monfort and Renault (1993)) and the Simulated Generalized Method of Moments (Duffie and Singleton (1993)).[2]

3.1. Moment conditions for the macroeconomic factors

To simulate the factor dynamics in Eq. (1), we rely on a Milstein approximation scheme, with discrete interval Δ, say. We simulate H paths of length T of the two observable factors, and

[2] The estimators we develop are not as efficient as Maximum Likelihood. Under some conditions, the methods put forward by Gallant and Tauchen (1996), Fermanian and Salanié (2004), Carrasco, Chernov, Florens and Ghysels (2007), Aït-Sahalia (2008), or Altissimo and Mele (2009), are asymptotic equivalent to Maximum Likelihood. In our context, they deliver asymptotically efficient estimators for the parameters in the first step. However, hinging upon these approaches in the remaining steps would make the two issues of unobservability of volatility and, especially, parameter estimation error considerably beyond the scope of this paper.

sample them at the same frequency as the available data, obtaining $y^{\phi}_{1,t,\Delta,h}$ and $y^{\phi}_{2,t,\Delta,h}$, where $y^{\phi}_{j,t,\Delta,h}$ is the value at time t taken by the j-th factor, at the h-th simulation performed with $\boldsymbol{\phi}$—the parameter vector relating to the process underlying the macroeconomic factors. Then, we estimate the following auxiliary models on both historical and simulated data, for $i = 1, 2$,[3]

$$y_{i,t} = a^y_i + \sum_{j\in\{12,24\}} b^y_{i,1,j} y_{1,t-j} + \sum_{j\in\{12,24\}} b^y_{i,2,j} y_{2,t-j} + \epsilon^y_{i,t}, \tag{17}$$

and

$$y^{\phi}_{i,t,\Delta,h} = a^y_{i,h} + \sum_{j\in\{12,24\}} b^y_{i,1,j,h} y^{\phi}_{1,t-j,\Delta,h} + \sum_{j\in\{12,24\}} b^y_{i,2,j,h} y^{\phi}_{2,t-j,\Delta,h} + \epsilon^y_{i,t,h}. \tag{18}$$

Next, let $\tilde{\boldsymbol{\varphi}}_T = \left(\tilde{\boldsymbol{\varphi}}_{1,T}, \tilde{\boldsymbol{\varphi}}_{2,T}, \bar{y}_1, \bar{y}_2, \hat{\sigma}_1, \hat{\sigma}_2\right)^\top$ where $\tilde{\boldsymbol{\varphi}}_{1,T}$ and $\tilde{\boldsymbol{\varphi}}_{2,T}$ denote the ordinary least squares (OLS, henceforth) estimators of the parameters in Eq. (17), and $\bar{y}_i$ and $\hat{\sigma}_i$ are the sample mean and standard deviation of the macroeconomic factors. Let $\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi})$ be the simulated counterpart to $\tilde{\boldsymbol{\varphi}}_T$ at simulation h, including the OLS estimator of the parameters in Eq. (18), and the sample means and standard deviations of the macroeconomic factors. The estimator of $\boldsymbol{\phi}$ is:

$$\hat{\boldsymbol{\phi}}_T \equiv \arg\min_{\boldsymbol{\phi}\in\boldsymbol{\Phi}_0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi}) - \tilde{\boldsymbol{\varphi}}_T \right\|^2, \tag{19}$$

where $\boldsymbol{\Phi}_0$ is some compact set. Appendix B in the Supplemental material develops the asymptotic theory relating to this estimator.

3.2. Moment conditions for realized returns and volatility

Data on macroeconomic factors and stock returns do not allow us to identify the structural parameters of the model. In particular, there are many combinations of $\boldsymbol{\delta} = (\delta_j)^3_{j=0}$ and $\boldsymbol{\lambda} = \left(\lambda_{1(j)}, \lambda_{2(j)}\right)^3_{j=1}$ in Eqs. (9)-(11), giving rise to the same asset price. In this second step, we estimate the parameters $\boldsymbol{\theta}_G = (g, \sigma_g)$ in Eq. (4), the reduced-form parameters, $(s_j)^3_{j=0}$ in Eqs. (9)-(11), and the parameters for the unobservable factor, $(\kappa_3, \mu_3, \alpha_3, \beta_3)$. The parameters $\boldsymbol{\lambda}$ shall be estimated in a third and final step, described in the next section. Note that, theoretically, it might be possible to collapse the second and third steps of our estimation procedure into a single

[3] The choice of lags for all the auxiliary models in Section 3 relies on the BIC criterion, and our additional concern to have non-overlapping regressors—with the exception of a lag 6 in Eq. (23), which was empirically important.

one, where a combined use of data on dividends and volatility derivatives might help identify $\boldsymbol{\delta}$ and $\boldsymbol{\lambda}$. We do not pursue this approach because it revealed to be computationally prohibitive.

Even proceeding in this way, we cannot tell apart the loading on the unobservable factor, s_3, from the parameters underlying the dynamics of this factor, $(\kappa_3, \mu_3, \alpha_3, \beta_3)$, as this is independent of the observable ones. We impose the normalization $\mu_3 \equiv 1$. We estimate $\boldsymbol{\theta}_G$ using the time-series of the low-frequency component of the real stock price growth, extracted through the Hodrick-Prescott filter with smoothing parameter equal to 14400, given we are using monthly data (Hodrick and Prescott (1997)). We simulate H paths of length T of the unobservable factor $y_3(t)$, and the secular growth, $G(t)$, using a Milstein approximation with discrete interval Δ, and sample them at the same frequency as the data, obtaining for $\boldsymbol{\theta}_u = (\kappa_3, \alpha_3, \beta_3, s_3)$ and $\hat{\boldsymbol{\theta}}_{G,T} = (\hat{g}_T, \hat{\sigma}_{G,T})$, and simulation h, the series $y_{3,t,\Delta,h}^{\theta_u}$ and $G_{t,\Delta,h}^{\hat{\theta}_{G,T}}$. Likewise, let $S_{t,\Delta,h}^{\theta}(\hat{\boldsymbol{\theta}}_{G,T})$ be the simulated series of the stock price, when the parameters are fixed at $\boldsymbol{\theta} = (\boldsymbol{\theta}_u, (s_j)_{j=0}^3)$ and $\hat{\boldsymbol{\theta}}_{G,T}$:

$$\ln S_{t,\Delta,h}^{\theta}(\hat{\boldsymbol{\theta}}_{G,T}) = \ln G_{t,\Delta,h}^{\hat{\theta}_{G,T}} + \ln\left(s_0 + s_1 y_{1,t} + s_2 y_{2,t} + s_3 y_{3,t,\Delta,h}^{\theta_u}\right), \tag{20}$$

where $G_{0,\Delta,h}^{\hat{\theta}_{G,T}} \equiv 1$, as in Eq. (4). We fix the intercept, s_0, so as to make the model-implied average of the detrended stock price match its empirical counterpart: $s_0 = \bar{S}^d - s_1\bar{y}_1 - s_2\bar{y}_2 - s_3$, where $\bar{S}^d$ denotes the sample mean of the detrended stock price $S_t^d \equiv e^{-\hat{g}_T t} S_t$, S_t is the real stock price index observed at time t, and finally, $\bar{y}_1$ and $\bar{y}_2$ are the sample means of the two macroeconomic factors $y_{1,t}$ and $y_{2,t}$. Note, we simulate the stock price using the *observed* samples of $y_{1,t}$ and $y_{2,t}$, a feature of the estimation strategy that results in improved efficiency, as discussed below.

Following Mele (2007) and Fornari and Mele (2010), we measure the volatility of the monthly continuously compounded price changes, as:

$$\mathrm{Vol}_t = \sqrt{6\pi} \cdot \frac{1}{12} \sum_{i=1}^{12} \left| \ln\left(\frac{S_{t+1-i}}{S_{t-i}}\right) \right|. \tag{21}$$

Next, define yearly returns as $R_t = \ln(S_t/S_{t-12})$, and let $R_{t,\Delta,h}^{\theta}(\hat{\boldsymbol{\theta}}_{G,T})$ and $\mathrm{Vol}_{t,\Delta,h}^{\theta}(\hat{\boldsymbol{\theta}}_{G,T})$ be the simulated counterparts to R_t and Vol_t.

Our estimator relies on two auxiliary models that capture the main statistical facts about stock returns and return volatility in our dataset. The auxiliary model for returns is:

$$R_t = a^{\mathrm{R}} + b_1^{\mathrm{R}} y_{1,t-12} + b_2^{\mathrm{R}} y_{2,t-12} + \epsilon_t^{\mathrm{R}}, \tag{22}$$

and that for return volatility is:

$$\mathrm{Vol}_t = a^{\mathrm{V}} + \sum_{i\in\{6,12,18,24,36,48\}} b_i^{\mathrm{V}} \mathrm{Vol}_{t-i} + \sum_{i\in\{12,24,36,48\}} b_{1,i}^{\mathrm{V}} y_{1,t-i} + \sum_{i\in\{12,24,36,48\}} b_{2,i}^{\mathrm{V}} y_{2,t-i} + \epsilon_t^{\mathrm{V}}. \quad (23)$$

Let $\tilde{\boldsymbol{\vartheta}}_T = \left(\tilde{\boldsymbol{\vartheta}}_{1,T}, \tilde{\boldsymbol{\vartheta}}_{2,T}, \bar{R}, \overline{\mathrm{Vol}}\right)^\top$, where $\bar{R}$ and $\overline{\mathrm{Vol}}$ are the sample means of stock returns and volatility, $\tilde{\boldsymbol{\vartheta}}_{1,T}$ is the OLS estimate of the parameters in Eq. (22) and $\tilde{\boldsymbol{\vartheta}}_{2,T}$ is the OLS estimate of the parameters in Eq. (23). Let $\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}_{G,T})$ be the simulated counterpart to $\tilde{\boldsymbol{\vartheta}}_T$ at simulation h, using $R_{t,\Delta,h}^{\theta}(\hat{\boldsymbol{\theta}}_{G,T})$ and $\mathrm{Vol}_{t,\Delta,h}^{\theta}(\hat{\boldsymbol{\theta}}_{G,T})$. The estimator of $\boldsymbol{\theta} = (\boldsymbol{\theta}_u, (s_j)_{j=0}^3)$ is:

$$\hat{\boldsymbol{\theta}}_T = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}_{G,T}) - \tilde{\boldsymbol{\vartheta}}_T \right\|^2, \quad (24)$$

where $\boldsymbol{\Theta}_0$ is a compact set. As shown in detail in Appendix B of the Supplemental material, the structure of the asymptotic covariance matrix of this estimator differs from that of $\hat{\boldsymbol{\phi}}_T$ in Eq. (19), due to two reasons. First, stock price paths are simulated through Eq. (20), with secular growth parameters fixed at their estimates, $\hat{\boldsymbol{\theta}}_{G,T}$, leading to parameter estimation error, which is asymptotically accounted for. Second, ours is, in fact, a conditional simulated inference estimator, in that the simulations in Eq. (20) occur conditionally upon the sample realizations of the observable factors, $y_{1,t}$ and $y_{2,t}$. This feature of the method results in a correlation among the auxiliary parameter estimates obtained over all the simulations, and leads to an efficiency improvement over unconditional (simulated) inference.

3.3. Estimation of the risk-premium parameters

We estimate the risk-premium parameters, $\boldsymbol{\lambda}$, by matching moments and impulse response functions of the model-based VIX, $\mathrm{VIX}\left(\boldsymbol{y}\left(t\right)\right)$ in Eq. (15), to those of the model-free VIX index, VIX_t in Eq. (14), with $t_0 - t$ equal to one month. Since the new VIX index is available only since 1990, we use a sample of $\mathcal{T}$ observations in this step, with $\mathcal{T} < T$. Whilst $\mathrm{VIX}\left(\boldsymbol{y}\left(t\right)\right)$ is not known in closed-form, it can be accurately approximated through simulations, as explained in Appendix B of the Supplemental material. Note, also, that in the actual computation of Eq. (15), we replace the unknown parameters, s_0, $\left(s_j, \kappa_j, \alpha_j, \beta_j\right)_{j=1}^3$, $(\bar{\kappa}_i, \mu_i)_{i=1}^2$ and σ_G, with their estimated counterparts computed in the previous two steps: $\hat{\boldsymbol{\theta}}_T$, $\hat{\boldsymbol{\phi}}_T$ and $\hat{\sigma}_{G,T}$. As in the previous step, we use the observed samples of the macroeconomic factors $y_{1,t}, y_{2,t}$, and simulate samples for the latent factor only. We rely on the following auxiliary model:

$$\mathrm{VIX}_t = a^{\mathrm{VIX}} + b^{\mathrm{VIX}}\mathrm{VIX}_{t-1} + \sum_{i\in\{36,48\}} b_{1,i}^{\mathrm{VIX}} y_{1,t-i} + \sum_{i\in\{36,48\}} b_{2,i}^{\mathrm{VIX}} y_{2,t-i} + \epsilon_t^{\mathrm{VIX}}. \tag{25}$$

Define, $\tilde{\boldsymbol{\psi}}_{\mathcal{T}} = \left(\tilde{\boldsymbol{\psi}}_{1,\mathcal{T}}, \overline{\mathrm{VIX}}, \hat{\sigma}_{\mathrm{VIX}}\right)^\top$, where $\tilde{\boldsymbol{\psi}}_{1,\mathcal{T}}$ is the OLS estimator of the parameters in Eq. (25), and $\overline{\mathrm{VIX}}$ and $\hat{\sigma}_{\mathrm{VIX}}$ are the sample mean and standard deviation of the VIX index. Likewise, define $\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda})$, the simulated counterpart to $\tilde{\boldsymbol{\psi}}_{\mathcal{T}}$ at simulation h, obtained through simulations of the model-implied index, $\mathrm{VIX}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda})$ say, where the paths of the two macroeconomic factors, $y_{1,t}$ and $y_{2,t}$, are fixed at their sample values. The estimator of $\boldsymbol{\lambda}$ is:

$$\hat{\boldsymbol{\lambda}}_{\mathcal{T}} = \arg\min_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}_0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}) - \tilde{\boldsymbol{\psi}}_{\mathcal{T}} \right\|^2, \tag{26}$$

for some compact set $\boldsymbol{\Lambda}_0$. This estimator is, similarly as $\hat{\boldsymbol{\theta}}_T$ in Eq. (24), affected by parameter estimation error, arising because $\mathrm{VIX}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda})$, the model-implied VIX index, is simulated using parameters estimated in the previous two steps, $\hat{\boldsymbol{\phi}}_T$, $\hat{\boldsymbol{\theta}}_T$ and $\hat{\sigma}_{G,T}$. At the same time, the estimator $\hat{\boldsymbol{\lambda}}_{\mathcal{T}}$ in Eq. (24) is a conditionally simulated one, in that it relies on the observations of the macroeconomic factors $y_{1,t}$ and $y_{2,t}$, thereby resulting in efficiency gains.

3.4. Bootstrap Standard Errors

The limiting variance-covariance matrices for $\hat{\boldsymbol{\phi}}_T$ in Eq. (19), $\hat{\boldsymbol{\theta}}_T$ in Eq. (24), and $\hat{\boldsymbol{\lambda}}_{\mathcal{T}}$ in Eq. (26) are characterized in Appendix B.1 of the Supplemental material. They are not known in closed form, and must be estimated through the computation of several numerical derivatives. Moreover, our sample sizes are relatively small, compared to those we usually have access to in empirical finance, and in particular such is that available for the estimation of the risk premium parameters. We rely on bootstrap standard errors consistent for those implied by the asymptotic variance-covariance matrices for $\hat{\boldsymbol{\phi}}_T$, $\hat{\boldsymbol{\theta}}_T$ and $\hat{\boldsymbol{\lambda}}_{\mathcal{T}}$. Bootstrap standard errors are not only easier to compute, but also less prone to numerical errors, and likely to be more accurate than those based on asymptotic approximations, in finite samples. Finally, the auxiliary models we utilize are potentially misspecified, and they likely lead to a score that is not a martingale difference sequence. We appeal to the "block-bootstrap" to address this technical issue. Appendix B.2 of the Supplemental material develops results and algorithms that allow us to make use of this method within the simulation-based estimation procedure of this section.

4. Empirical analysis

4.1. Data

Our security data include the S&P 500 Compounded index, and the VIX index maintained by the CBOE. The VIX index is available daily, but only after January 1990. Our macroeconomic variables include the consumer price index (CPI), and the seasonally adjusted industrial production (IP) index for the US. Information related to the CPI and the IP indexes is made available to the market between the 19-th and the 23-th of every month. To possibly avoid overreaction to releases of information, we sample the S&P Compounded index and the VIX index every 25-th of the month. We compute the real stock price as the ratio between the S&P index and the CPI. Our dataset, then, includes (i) monthly observations of the VIX index, from January 1990 to December 2006, for a total of 204 observations; and (ii) monthly observations of the real stock price, the CPI and the IP indexes, from January 1950 to December 2006, for a total of 672 observations.

Our dataset also includes monthly observations of the University of Michigan Consumer Sentiment index, from January 1978 to December 2006 (for a total of 336 observations), Finally, we utilize additional data, from January 2007 to March 2009, to implement a stress test of how the previously estimated model would have performed over a particularly critical period. This out-of-sample period is critical for at least three reasons: first, the NBER determined that the US economy entered in a recession in December 2007, which is the third NBER-dated recession since the creation of the new VIX index; second, this period includes the quite unique events leading to the subprime crisis; third, both realized stock market volatility and the VIX index reached record highs, and possibly pose challenges to rational models of asset prices. Our out-of-sample experiments are not intended to forecast the market, stock market volatility, and the level of the VIX index. Rather, we feed the model estimated up to December 2006, with macroeconomic data (the CPI and IP indexes) available from January 2007, and compare the predictions of the model with the actual movements of the market, stock market volatility and the VIX index.

Many theoretical explanations and, in fact, the empirical evidence, would lead us to expect that asset prices are, indeed, related to variables tracking business cycles (see, e.g., Cochrane (2005)), such as the CPI and the IP growth. For example, in their seminal article relating stock returns to the macroeconomy, Chen, Roll and Ross (1986) find that industrial production growth and inflation are among the most prominent priced factors. Theoretically, in standard theories of external habit formation, the pricing kernel volatility is driven by the surplus consumption ratio, defined as the percentage deviation of current consumption, C, from some habit level, H, i.e. $(C-H)/C$, which highly correlates with procyclical variables such as industrial production growth. Likewise, standard asset pricing models predict that compensation for inflation risk relates to variables that are highly correlated with inflation (e.g., Bakshi and Chen (1996), Buraschi and Jiltsov (2005)). Mainly for computational reasons, we refrain from considering additional factors to model the linkages of the pricing kernel to the business cycle.

Figure 1 depicts the two series $y_{1,t}$ (year-to-year gross inflation) and $y_{2,t}$ (year-to-year industrial production growth) along with NBER-dated recession events. Gross inflation is procyclical, although it peaked up during the 1975 and the 1980 recessions, as a result of the geopolitical driven oil crises that occurred in 1973 and 1979. Its volatility during the 1970s was large until the Monetary experiment of the early 1980s, although it dramatically dropped during the period following the experiment, usually referred to as the Great Moderation (e.g., Bernanke (2004)). At the same time, inflation is persistent: a Dickey-Fuller test rejects the null hypothesis of a unit root in $y_{1,t}$, although the rejection is at the marginal 95% level. The inclusion of inflation as a determinant of the pricing kernel displays one attractive feature. An old debate exists upon whether stocks provide a hedge against inflation (see, e.g., Danthine and Donaldson (1986)). While our no-arbitrage model is silent about the general equilibrium forces underlying inflation-hedge properties of asset prices, its data-driven structure allows us to assess quite directly the relations between inflation and the stock price, returns, volatility and volatility risk-premiums.

Figure 1 also shows that while the volatility of industrial production growth dropped during the

Great Moderation, growth is still persistent, although less so than gross inflation: here, a Dickey-Fuller test rejects the null hypothesis of a unit root in $y_{2,t}$ at any conventional level. Finally, the properties of inflation and industrial production growth over our out-of-sample period, from January 2007 to March 2009, are discussed in Section 4.2.4.

4.2. Estimation results

4.2.1. Macroeconomic drivers

Table 1 reports parameter estimates and block-bootstrap standard errors for the joint process of the two macroeconomic variables, $y_{1,t}$ and $y_{2,t}$, as set forth in Section 3.1. The estimates are all largely significant, and confirm our discussion of Figure 1: inflation is more persistent than IP growth, as both its speed of adjustment in the absence of feedbacks, κ_1, and its feedback parameter, $\bar{\kappa}_1$, are much lower than the counterparts for IP growth, κ_2 and $\bar{\kappa}_2$. Finally, the estimates of β_1 and β_2 are both negative, implying that the volatility of these two macroeconomic variables are countercyclical, an interesting property, from an asset pricing perspective. However, we note that the estimate of β_1, albeit statistically significant, is also economically very small.[4]

4.2.2. Aggregate stock returns and volatility

Table 2 reports estimates and block-bootstrap standard errors for (i) the parameters affecting secular growth, (ii) the parameters linking the two macroeconomic factors and the unobservable factor to the asset price, and (iii) the parameters for the unobservable factor process, as explained in Section 3.1. The estimates are all largely significant and point to two conclusions. First, the stock price is positively related to IP growth and negatively related to inflation. Second, the unobservable factor is quite persistent, displaying high volatility, as the estimate of the speed reversion coefficient, κ_3 is low. Note, the literature on long run risks started by Bansal and Yaron (2004)

[4] It is known since at least Friedman (1977) that high variability of inflation might link to high inflation. For example, Engle (1982) finds that inflation volatility increases during the middle 1970s. We have constructed measures of inflation volatility similar to that in Eq. (21), relating to the first difference of inflation, which confirm these findings. We also find that after the 1970s, inflation slowdowns tend to occur more rapidly than inflation increases although overall, a clear relation between inflation and inflation volatility is hard to establish. The estimate of β_1 for our continuous time model is likely to reflect these facts.

emphasizes the asset pricing implications of long-run risks affecting the expected consumption growth rate. Interestingly, the presence of a persistent factor affecting stock returns and volatility emerges quite neatly from our estimation. Note, however, that in long-run risk models, expected consumption growth is unlikely to affect the dynamics of stock volatility which, instead, are inherited by those of the volatility of consumption growth. In our model, our unobservable factor does, instead, affect stock volatility, and substantially, as explained below.

Figure 2 shows the dynamics of stock returns and volatility predicted by the model, along with their sample counterparts, calculated as described in Section 3.2. These predictions are obtained by feeding the model with sample data for the two macroeconomic factors, $y_{1,t}$ and $y_{2,t}$, in conjunction with simulations of the third unobservable factor, using all the estimated parameters. For each point in time, we average over the cross-section of 1000 simulations, and report returns (in the top panel of Figure 2) and volatility (in the bottom panel). Returns are computed as we do with the data, and volatility is obtained through Eq. (12).

The model appears to capture the procyclical nature of stock returns and the countercyclical behavior of stock volatility. It generates *all* the stock market drops occurred during the NBER recessions, and *all* the volatility upward swings occurred during the NBER recessions, including the dramatic spike of the 1975 recession. In the data, average stock volatility is about 11.50%, with a standard deviation of about 4.0%. The model predicts an average volatility of about 13%, with a standard deviation of about 3.1%.

How much of the variation in volatility can be attributable to macroeconomic factors? It is a natural question, as the key innovation of our model is the introduction of these factors for the purpose of explaining volatility, on top of a standard unobservable factor. We address this issue and calculate: (i) the ratio of the instantaneous return variance due to factor y_j, $s_j^2\left(\alpha_j+\beta_j y_j(t)\right)$, to the total instantaneous variance, $\sigma^2(t)$ in Eq. (12), as well as (ii) the ratio of the instantaneous variance of secular growth, σ_G^2, to $\sigma^2(t)$, as follows,

$$C_j(t) \equiv \frac{s_j^2\left(\alpha_j+\beta_j y_j(t)\right)}{\sigma^2(t)},\ j=1,2,3, \quad \text{and} \quad C_G(t) \equiv \frac{\sigma_G^2}{\sigma^2(t)}. \tag{27}$$

Figure 3 depicts the time series of $C_j(t)$ and $C_G(t)$ implied by our estimated model, obtained, as usual, by feeding the model with the observed samples of $y_1(t)$ and $y_2(t)$, and averaging across

1000 dynamic simulations of $y_3(t)$. The clear finding is that industrial production growth makes the most important contribution to stock volatility: the time series average of $C_2(t)$ is above 73%, more than four times higher than $C_3(t)$, the contribution made by the unobserved factor. Panel A of Table 4 reports averages and standard deviations of the contributions made by all factors, and secular growth. Variations in industrial production growth and the unobserved factor are responsible, alone, for more than 90% of the variation in stock volatility. It is a striking result, as one challenge we face is to explain why we have observed a sustained stock market volatility, in spite of the Great Moderation. Our estimated model entails two clear conclusions.

First, as Figure 3 makes clear, the 73% average contribution of industrial production growth to stock volatility seems to be rather stable over time, at least once we exclude the 1950s—a period of sustained volatility for growth (see Figure 1). Accordingly, the Great Moderation does merely appear to have affected the variability of the linkages between industrial production growth and aggregate stock volatility, not the very same linkages. To illustrate, Panel A of Table 4 shows averages and standard deviations of the factors' contributions across different sampling periods. We take 1982 to be the year that marks the beginning of the Great Moderation, characterized by the inauguration of the Federal Reserve monetary policy turning point and a lower volatility of real macroeconomic variables (e.g., Blanchard and Simon (2001)). As is clear, whilst the average contributions are stable, the variability of these contributions has decreased over the Great Moderation. For example, the average of $C_2(t)$ is between 73% - 75%, across all sampling periods, whereas its standard deviation decreases to 3.47% during the 1982-2006 sample, from 9.65% (1950-1981) and 5.03% (1960-1981).

Second, the contribution of industrial production growth to volatility, albeit crucial, is not exhaustive. Our model predicts that stock volatility cannot be explained by macroeconomic variables only, as the unobserved factor accounts for about 17% of the fluctuations in $\sigma^2(t)$. Equally important is the observation that the contribution of industrial production to stock volatility is strongly countercyclical, exhibiting large upward swings starting at, and sometimes, anticipating, turning points, as in the case of the 1970s recessions and the most recent, 2001 recession. Instead, the contribution of the unobserved factor to stock volatility, $C_3(t)$, is procyclical, for the simple reason that the instantaneous volatility of $y_3(t)$ does not obviously link to the business cycle, thereby making the ratio $C_3(t)$ in Eq. (27) procyclical, due to the countercyclical nature of its denominator, $\sigma^2(t)$. All in all, our empirical results suggest that while unobserved factors are needed to explain the level of stock volatility, industrial production is needed to explain the countercyclical swings of stock volatility that we have in the data—the volatility of volatility.

Finally, the contribution of secular growth to stock volatility is limited, being approximately 8%, and that of gross inflation plays an even more marginal role, being less than 1%. Note, however, that our model predicts that inflation links to asset returns and volatility in a manner comparable to that in the data. For example, it is well-known since at least Fama (1981) that real stock returns are negatively correlated with inflation, a property that hinders the ability of stocks to hedge against inflation. In our sample, this correlation is -35%, while the correlation our model generates is -24%. Finally, the correlation between stock volatility and inflation is about 20% in the data, while that implied by the model is about 25%.

The predictions of the model discussed so far rely on cross-sectional averages of dynamic simulations of the unobserved factor, y_3. Yet what is the interpretation of this unobserved factor? Let us invert the price function in Eq. (8), for y_3, and for each month, as follows:

$$-\hat{y}_{3,t} \equiv -\frac{1}{\hat{s}_3}\left(\frac{s_t}{\hat{G}_t} - \hat{s}_0 - \hat{s}_1 y_{1,t} - \hat{s}_2 y_{2,t}\right), \tag{28}$$

where $(\hat{s}_j)_{j=0}^3$ are estimates of the pricing function coefficients, as reported in Table 2, and $\hat{G}_t$ is the cross-sectional average of 1000 dynamic simulations of secular growth.

Figure 4 (top panel) depicts $-\hat{y}_{3,t}$ (in bold), along with 100 simulated trajectories of the unobserved factor performed with the parameter estimates in Table 2. Reinsuringly, the range of variation of the model-implied factor roughly falls within that of the simulated trajectories of this factor. Note that the estimate of s_3 is negative, such that $-\hat{y}_{3,t}$ positively affects the real stock price—higher realizations of $-\hat{y}_{3,t}$ amount to good pieces of news to the stock market. There are episodes where $-\hat{y}_{3,t}$ comes close to the edges of the realized range of variation experienced by the unobserved factors during the simulations. These episodes are interesting, as they correspond to: (i) the lows of the late 1970s and the early 1980s, and (ii) the highs of the dotcom bubble that occurs in the late 1990s. The extracted factor oscillates between (about) its minimum and its maximum over those approximate twenty years. The rise and fall of this period have, then, a clear economic interpretation, with the late 1970s and early 1980s being particularly bad times, marked by the occurrence of a double dip recession, and the extraordinary market boom over the dotcom bubble being notoriously suspected to be one of exuberance (e.g., Shiller (2005)). These observations motivate us to explore the extent to which our extracted factor links to indexes of "sentiment," following a recent strand of the literature that attempts to link asset price movements to factors such as investors uncertainty (as in David and Veronesi (2006)), confidence risk (as in Bansal and Shaliastovich (2010)), or Knightian uncertainty (as in Drechsler (2010) or Mele and Sangiorgi (2011)). The bottom panel of Figure 4 compares the time series behavior of the model-implied unobserved factor, $-\hat{y}_{3,t}$, with that of an index of consumer confidence—the University of Michigan Consumer Sentiment (UMCSENT) index, available from January 1978.

Note how the UMCSENT index tracks the lows and the highs of the market that have so slowly occurred over the last thirty years: the bad times of the late 1970s and early 1980s, the rise occurring over the late 1980s and culminating with the dotcom bubble of the late 1990s and, finally, the drop of the late 2000s, corresponding to the subprime crisis—a period we study in detail in the next section. Interestingly, our extracted factor, $-\hat{y}_{3,t}$, co-moves positively with the UMCSENT index, correlating with it at about 50%. In contrast, its correlation with the macroeconomic factors is modest (10% with inflation and 30% with industrial production growth). Interestingly, then, the pattern our extracted factor exhibits is one that mostly tracks long-run movements of the market, even more so than the short-term movements relating to business cycles.

4.2.3. Volatility risk-premiums and the dynamics of the VIX index

Table 3 reports parameter estimates and block-bootstrap standard errors for the vector of the risk-premiums coefficients $\boldsymbol{\lambda}$ in Eq. (7), as set forth in Section 3.3. The estimates, all significant, imply that the risk-premiums processes are all positive, and quite large, especially those relating to the two macroeconomic factors. Moreover, the risk compensation for inflation increases with inflation and that for industrial production is countercyclical, given the sign of the estimated values for the loadings of inflation, $\left(\beta_1\lambda_{1(1)} + \lambda_{2(1)}\right)$ (positive), and industrial production,

$\left(\beta_2\lambda_{1(2)} + \lambda_{2(2)}\right)$ (negative), in the risk-premium process of Eq. (7). While gross inflation does receive compensation, the countercyclical behavior of the risk-premium for industrial production growth is even more critical, as we explain below. Our estimated model predicts that in bad times, the risk-premium for industrial production growth goes up, and future expected economic conditions even worsen, under the risk-neutral probability, which boosts future expected volatility, under the same risk-neutral probability. In part because of these effects, the VIX index predicted by the model is countercyclical. This reasoning is quantitatively sound. Figure 5 (top panel) depicts the VIX index, along with the VIX index predicted by the model and the (square root of the) model-implied expected integrated variance. The model appears to track the swings the VIX index has undergone over the 1991 and the 2001 recession episodes.

The top panel of Figure 5 also shows the dynamics of volatility expected under the physical probability. This expected volatility is certainly countercyclical, although it does not display the large variations the model predicts for its risk-neutral counterpart, the VIX index. The VIX index predicted by the model is countercyclical because, as explained, the risk-premiums required to bear the fluctuations of the macroeconomic factors are (i) positive and (ii) countercyclical, and, also, because (iii) current volatility is countercyclical. Under the physical probability, expected volatility is countercyclical only because of the third effect. However, quantitatively, movements of volatility risk-premiums account for variations in the VIX index sensibly more than those of the volatility expected under the physical probability, as clearly summarized by Figure 5.

Which factors mostly contribute to the dynamics of the VIX? Panel B of Table 4 reports averages and standard deviations of the contributions of each factor, as predicted by our estimated model. We calculate each of these contributions by evaluating C_j and C_G in Eq. (27) under the risk-neutral probability and, then, aggregating the average paths of C_j and C_G for every month, and, finally, taking cross-sectional averages over 1000 dynamic simulations of the unobserved factor. Similarly as for the results in Section 4.2.2 on realized volatility, we find, again, our model predicts industrial production growth to account for the bulk of variation of the VIX index. The unobserved factor accounts for less than 10%, and inflation and secular growth play a quite marginal role, explaining no more than 5%, of the model-implied VIX. Interestingly, Stock and

Watson (2003) find that the linkages of asset prices to growth are stronger than for inflation. Our results further qualify this finding: inflation does not seem to affect too much the dynamics of neither realized volatility nor future expected volatility under the risk-neutral probability.

Finally, the bottom panel in Figure 5 plots the volatility risk-premium, defined as in Eq. (16). This risk-premium is countercyclical, and this property arises for exactly the same reasons we put forward to explain the swings the model predicts for the VIX index: positive compensation for risk, combined with countercyclical variation of the premiums required to compensate for the risk in fluctuations of the macroeconomic factors.

4.2.4. Out-of-sample predictions of the model, and the subprime crisis

We undertake out-of-sample experiments to investigate the model's predictions over a quite exceptional period, that from January 2007 to March 2009. This sample covers the subprime turmoil, and features unprecedented events, both for the severity of capital markets uncertainty and the performance of the US economy. The market witnessed to a spectacular drop accompanied by an extraordinary surge in volatility. In March 2009, yearly returns plummeted to -58.30%, a performance even worse than that experienced in October 1974 (-58.10%). Furthermore, according to our estimates, obtained through Eq. (21), aggregate stock volatility reached 28.20% in March 2009, the highest level ever experienced in our sample. Finally, the VIX index hit its highest value in our sample in November 2008 (72.67%), and remained stubbornly high for several months. The time series behavior of stock returns, stock volatility and the VIX index during our out-of-sample period are depicted over the shaded areas in Figures 2 and 5.

Macroeconomic developments over our out-of-sample period (the shaded area in Figure 1) were equally extreme, with yearly inflation rates achieving negative values in 2009, and yearly industrial production growth being as low as -13%, in March 2009. Under such macroeconomic conditions, we expect our model to produce the following predictions: (i) stock returns drop, (ii) stock volatility rises, (iii) the VIX index rises, and more than stock volatility. The mechanism is, by now, clear. Asset prices and, hence, returns, plummet, as they are positively related to inflation and growth, which both crashed. Moreover, volatility increases, with the VIX index increasing even more, due to our previous finding of (i) sizeable macroeconomic risk-premiums and (ii) strong countercyclical variation in these premiums.

Figures 2 and 5 confirm our reasoning, and reveal that the model is able to trace out the dynamics of stock returns and volatility (Figure 2), and the VIX index (Figure 5), over the out-of-sample period. The market literally crashes, as in the data, although only less than a half as much as in the data: the lowest value for yearly stock returns the model predicts, out-of-sample, is -21.77%, which is the second lowest figure our model produces, since after the quite volatile periods occurring over the 1950s and the early 1960s. (The lowest level the model predicts after those periods is -29.91%, for March 1975, the last month of the second severe recession of the 1970s.) Instead, the model predicts that stock volatility and the VIX index surge even more than in the data, reaching record highs of 26.68% (volatility) and 61.27% (VIX).

Figure 6 provide additional details about the period from January 2000 to March 2009. It compares stock volatility and the VIX index with the predictions of the model and those of a OLS regression. The OLS for volatility is that in Eq. (23), excluding the lag for six months, related to the autoregressive term. The OLS for the VIX index is that in Eq. (25). OLS predictions are obtained by feeding the OLS predictive part with its regressors, using parameter estimates obtained with data up to December 2006. The following table reports Root Mean Squared Errors (RMSE) for both our model and OLS, calculated over the out-of-sample period.

RMSE for the model and OLS

	Model	OLS
Volatility	0.0478	0.0700
VIX Index	0.1119	0.1319

Overall, OLS predictions do not seem to capture the countercyclical behavior of stock volatility. As for the VIX index, the OLS model (in fact, by Eq. (25), an autoregressive, distributed lag model) produces predictions that are not as accurate as the model, and generate overfit. The model, instead, predicts the swings we see in the data, in both the last two recession episodes. The RMSEs clearly favour the model against OLS, although it appears to do so more with realized volatility than with the VIX index, as Figure 6 informally reveals.

5. Conclusion

How does aggregate stock market volatility relate to the business cycle? This old question has been formulated at least since Officer (1973) and Schwert (1989a,b). We learnt from recent theoretical explanations that the countercyclical behavior of stock volatility can be understood as the result of a rational valuation process. However, how much of this countercyclical behavior is responsible for the sustained level aggregate volatility has experienced for centuries? This paper develops a model where approximately 75% of the variations in stock volatility can be explained by macroeconomic factors, and where some unobserved component is also needed to make stock volatility consistent with rational asset valuation.

We show that risk-premiums arising from fluctuations in this volatility are strongly countercyclical, certainly more so than stock volatility alone. In fact, the risk-compensation for the fluctuation of the macroeconomic factors is large and countercyclical, and helps explain the swings in the VIX index that we observe during recessions. We undertake out-of-sample experiments that cover the 2007-2009 subprime crisis, when the VIX reached a record high of more than 70%, which our model can at least partially track, through a countercyclical variation in the volatility risk-premiums. Again, our model predicts that a business cycle factor such as industrial production growth can explain more than 85% of the variations of the VIX index.

The key aspect of our model is that the relations among the market, stock volatility, volatility risk-premiums and the macroeconomic factors, are consistent with no-arbitrage. In particular, volatility is endogenous in our framework: the same variables driving the payoff process and the volatility of the pricing kernel, and hence, the asset price, are those that drive stock volatility and volatility-related risk-premiums. A question for future research is to explore whether the no-arbitrage framework in this paper can be used to improve forecasts of real economic activity. In fact, stock volatility and volatility risk-premiums are driven by business cycle factors, as this paper clearly demonstrates. A challenging and fundamental question is to explore the extent to which business cycle, stock volatility and volatility risk-premiums do endogenously develop.

References

Adrian, T. and J. Rosenberg, 2008. "Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk," *Journal of Finance* 63, 2997-3030.

Aït-Sahalia, Y., 2008. "Closed-Form Likelihood Expansions for Multivariate Diffusions," *Annals of Statistics* 36, 906-937.

Altissimo, F. and A. Mele, 2009. "Simulated Nonparametric Estimation of Dynamic Models," *Review of Economic Studies* 76, 413-450.

Ang, A. and M. Piazzesi, 2003. "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics* 50, 745-787.

Ang, A., M. Piazzesi and M. Wei, 2006. "What Does the Yield Curve Tell Us about GDP Growth?" *Journal of Econometrics* 131, 359-403.

Bakshi, G. S. and Z. Chen, 1996. "Inflation, Asset Prices, and the Term Structure of Interest Rates in Monetary Economies," *Review of Financial Studies* 9, 241-275.

Bakshi, G. and D. Madan, 2000. "Spanning and Derivative Security Evaluation," *Journal of Financial Economics* 55, 205-238.

Bakshi, G. and D. Madan, 2006. "A Theory of Volatility Spreads," *Management Science* 52, 1945-1956.

Bansal, R. and A. Yaron, 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance* 59, 1481-1509.

Bansal, R. and I. Shaliastovich, 2010. "Confidence Risk and Asset Prices," *American Economic Review P&P* 100, 537-541.

Bernanke, B. S., 2004. "The Great Moderation," *Remarks by the Federal Reserve Board Governor* at the meetings of the Eastern Economic Association, Washington, DC, February 20.

Blanchard, O. and J. Simon, 2001. "The Long and Large Decline in U.S. Output Volatility," *Brookings Papers on Economic Activity* 1, 135-64.

Bloom, N., 2009, "The Impact of Uncertainty Shocks," *Econometrica* 77, 623-685.

Bloom, N., M. Floetotto, and N. Jaimovich, 2009, "Really Uncertain Business Cycles," Working paper, Stanford University.

Bollerslev, T. and H. Zhou, 2006. "Volatility Puzzles: A Simple Framework for Gauging Return-Volatility Regressions," *Journal of Econometrics* 131, 123-150.

Bollerslev, T., M. Gibson and H. Zhou, 2011. "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option Implied and Realized Volatilities," *Journal of Econometrics* 160, 102-118.

Bollerslev, T., G. Tauchen and H. Zhou, 2009. "Expected Stock Returns and Variance Risk Premia," *Review of Financial Studies* 22, 4463-4492.

Brandt, M. W. and Kang, Q., 2004. "On the Relationship Between the Conditional Mean and Volatility of Stock Returns: a Latent VAR Approach," *Journal of Financial Economics* 72, 217-257.

Britten-Jones, M. and A. Neuberger, 2000. "Option Prices, Implied Price Processes and Stochastic Volatility," *Journal of Finance* 55, 839-866.

Buraschi A. and A. Jiltsov, 2005. "Inflation Risk Premia and the Expectations Hypothesis," *Journal of Financial Economics* 75, 429-490.

Campbell, J. Y., 2003. "Consumption-Based Asset Pricing," in Constantinides, G.M., M. Harris and R. M. Stulz (Editors): *Handbook of the Economics of Finance* (Volume 1B, chapter 13), North-Holland Elsevier, 803-887.

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, 205-251.

Campbell, J. Y., M. Lettau, B. G. Malkiel and Y. Xu, 2001. "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance* 56, 1-43.

Carr, P. and D. Madan, 2001. "Optimal Positioning in Derivative Securities," *Quantitative Finance* 1, 19-37.

Carr, P. and L. Wu, 2009. "Variance Risk Premiums," *Review of Financial Studies* 22, 1311-1341.

Carrasco, M., M. Chernov, J.-P. Florens and E. Ghysels, 2007. "Efficient Estimation of General Dynamic Models with a Continuum of Moment Conditions," *Journal of Econometrics* 140, 529-573.

Chen, N.-F., R. Roll and S. A. Ross, 1986. "Economic Forces and the Stock Market," *Journal of Business* 59, 383-403.

Chernov, M. and E. Ghysels, 2000. "A Study towards a Unified Approach to the Joint Estimation of Objective and Risk-Neutral Measures for the Purpose of Options Valuation," *Journal of Financial Economics* 56, 407-458.

Cochrane, J. H., 2005. *Financial Markets and the Real Economy.* Foundations and Trends® in Finance. Hanover, Ma: now Publishers Inc.

Corradi, V. and W. Distaso, 2006. "Semiparametric Comparison of Stochastic Volatility Models using Realized Measures," *Review of Economic Studies* 73, 635-667.

Danthine, J.-P. and J. B. Donaldson, 1986. "Inflation and Asset Prices in an Exchange Economy," *Econometrica* 54, 585-605.

David, A. and P. Veronesi, 2006. "Inflation and Earnings Uncertainty and Volatility Forecasts," Working paper, University of Calgary and University of Chicago.

Drechsler, I., 2010. "Uncertainty, Time-Varying Fear, and Asset Prices," Working paper, New York University.

Drechsler, I. and A. Yaron, 2011. "What's Vol Got to Do With It," *Review of Financial Studies* 24, 1-45.

Duffee, G. R., 2002. "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance* 57, 405-443.

Duffie, D. and R. Kan, 1996. "A Yield-Factor Model of Interest Rates," *Mathematical Finance* 6, 379-406.

Dumas, B., 1995. "The Meaning of the Implicit Volatility Function in Case of Stochastic Volatility," Working paper, HEC Lausanne. Available at:
`http://www.insead.edu/facultyresearch/faculty/personal/bdumas/research/index.cfm`.

Engle, R. F., 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of the United Kingdom Inflation," *Econometrica* 50, 987-1007.

Engle, R. F. and J. G. Rangel, 2008. "The Spline GARCH Model for Low Frequency Volatility and Its Macroeconomic Causes," *Review of Financial Studies* 21, 1187-1222.

Fama, E. F., 1981. "Stock Returns, Real Activity, Inflation, and Money," *American Economic Review* 71, 545-565.

Fermanian, J.-D. and B. Salanié, 2004. "A Nonparametric Simulated Maximum Likelihood Estimation Method," *Econometric Theory* 20, 701-734.

Fornari, F. and A. Mele, 2010. "Financial Volatility and Economic Activity," Working paper, European Central Bank and London School of Economics.

Friedman, M., 1977. "Nobel Lecture: Inflation and Unemployment," *Journal of Political Economy* 85, 451-472.

Gallant, A. R. and G. Tauchen, 1996. "Which Moments to Match?" *Econometric Theory* 12, 657-681.

Garcia, R., M.-A. Lewis, S. Pastorello and E. Renault, 2011. "Estimation of Objective and Risk-Neutral Distributions Based on Moments of Integrated Volatility," *Journal of Econometrics* 160, 22-32.

Gouriéroux, C., A. Monfort and E. Renault, 1993. "Indirect Inference," *Journal of Applied Econometrics* 8, S85-S118.

Guvenen, F. 2009. "A Parsimonious Macroeconomic Model for Asset Pricing," *Econometrica* 77, 1711-1740.

Hamilton, J. D. and G. Lin, 1996. "Stock Market Volatility and the Business Cycle," *Journal of Applied Econometrics* 11, 573-593.

Heston, S. L., 1993. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies* 6, 327-343.

Hodrick, R. J. and E. C. Prescott, 1997. "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking* 29, 1-16.

King, M., E. Sentana and S. Wadhwani, 1994. "Volatility and Links between National Stock Markets," *Econometrica* 62, 901-933.

Mehra, R. and E. C. Prescott, 2003. "The Equity Premium in Retrospect," in Constantinides, G.M., M. Harris and R. M. Stulz (Editors): *Handbook of the Economics of Finance* (Volume 1B, chapter 14), North-Holland Elsevier, 889-938.

Mele, A., 2007. "Asymmetric Stock Market Volatility and the Cyclical Behavior of Expected Returns," *Journal of Financial Economics* 86, 446-478.

Mele, A. and F. Sangiorgi, 2011. "Ambiguity, Information Acquisition and Price Swings in Asset Markets," Working paper, London School of Economics and Stockholm School of Economics.

Schwert, G. W., 1989a. "Why Does Stock Market Volatility Change over Time?" *Journal of Finance* 44, 1115-1153.

Schwert, G. W., 1989b. "Business Cycles, Financial Crises and Stock Volatility," *Carnegie-Rochester Conference Series on Public Policy* 31, 83-125.

Shiller, R. J., 2005. Irrational Exuberance," Princeton: Princeton University Press.

Stock, J. H. and M. W. Watson, 2003. "Forecasting Output and Inflation: The Role of Asset Prices," *Journal of Economic Literature* 41, 788-829.

Tauchen, G., 2005. "Stochastic Volatility in General Equilibrium," Working paper, Duke University.

Todorov, V., 2010. "Variance Risk Premium Dynamics: The Role of Jumps," *Review of Financial Studies* 23, 345-383.

Todorov, V. and G. Tauchen, 2011. "Volatility Jumps," *Journal of Business and Economic Statistics* 29, 356-371.

Tables

Table 1

Parameter estimates and block-bootstrap standard errors for the joint process of the two macroeconomic factors, gross inflation, $y_{1,t} \equiv \mathrm{CPI}_t / \mathrm{CPI}_{t-12} \equiv y_1(t)$ and gross industrial production growth, $y_{2,t} \equiv \mathrm{IP}_t / \mathrm{IP}_{t-12} \equiv y_2(t)$, where CPI_t is the Consumer price index as of month t, IP_t is the real, seasonally adjusted industrial production index as of month t, and:

$$\begin{bmatrix} \mathrm{d}y_1(t) \\ \mathrm{d}y_2(t) \end{bmatrix} = \begin{bmatrix} \kappa_1 & \bar{\kappa}_1 \\ \bar{\kappa}_2 & \kappa_2 \end{bmatrix} \begin{bmatrix} \mu_1 - y_1(t) \\ \mu_2 - y_2(t) \end{bmatrix} \mathrm{d}t + \begin{bmatrix} \sqrt{\alpha_1 + \beta_1 y_1(t)} & 0 \\ 0 & \sqrt{\alpha_2 + \beta_2 y_2(t)} \end{bmatrix} \begin{bmatrix} \mathrm{d}W_1(t) \\ \mathrm{d}W_2(t) \end{bmatrix},$$

where $W_j(t)$, $j = 1, 2$, are two independent Brownian motions, and the parameter vector is $\boldsymbol{\phi}^\top = (\kappa_j, \mu_j, \alpha_j, \beta_j, \bar{\kappa}_j, j = 1, 2)$. Parameter estimates are obtained through the first step of the estimation procedure set forth in Section 3.1, relying on Indirect Inference and Simulated Method of Moments. Matching conditions relate to (i) parameter estimates for the auxiliary Vector Autoregressive models in Eq. (17), and (ii) the sample mean and standard deviation of $y_{1,t}$ and $y_{2,t}$. The sample covers monthly data for the period from January 1950 to December 2006.

	Estimate	Std error
κ_1	0.0231	0.0095
μ_1	1.0375	0.3784
α_1	$2.4408 \cdot 10^{-4}$	$1.0918 \cdot 10^{-4}$
β_1	$-1.0005 \cdot 10^{-6}$	$0.3374 \cdot 10^{-6}$
κ_2	0.9025	0.4037
μ_2	1.0386	0.3962
α_2	0.0253	0.0126
β_2	-0.0198	0.0084
$\bar{\kappa}_1$	-0.2995	0.1293
$\bar{\kappa}_2$	1.3723	0.6423

Table 2

Parameter estimates and block-bootstrap standard errors for stochastic secular growth, the real stock price and the unobservable factor:

$$S(t) = G(t)\left(s_0 + \sum_{i=1}^{3} s_i y_i(t)\right), \quad \frac{dG(t)}{G(t)} = g dt + \sigma_G dW_G(t),$$

where $S(t)$ is the real stock price, $G(t)$ is stochastic secular growth, W_G is a standard Brownian motion, $y_1(t)$ and $y_2(t)$ are the observed gross inflation and gross industrial production growth, as defined in Table 1, $y_3(t)$ is an unobserved factor, with the following dynamics:

$$\mathrm{d}y_3(t) = \kappa_3(\mu_3 - y_3(t))\,\mathrm{d}t + \sqrt{\alpha_3 + \beta_3 y_3(t)}\,\mathrm{d}W_3(t),$$

and $W_3(t)$ a standard Brownian motion. The parameter vector to be estimated is $\boldsymbol{\theta}^\top = (g, \sigma_G, \kappa_3, \mu_3, \alpha_3, \beta_3, (\delta_j, s_j)_{j=0}^{3})$, where the long run mean for the unobservable factor, μ_3, is set equal to one for the purpose of model's identification. Parameter estimates are obtained through the second step of the estimation procedure set forth in Section 3.2, relying on Indirect Inference and Simulated Method of Moments, with parameters (g, σ_G^2) estimated on the low frequency component of secular growth of the real stock price, extracted through the Hodrick-Prescott filter with smoothing parameter 1600. Matching conditions relate to (i) parameter estimates for the auxiliary model for stock returns, Eq. (22), and for the auxiliary model for stock volatility, Eq. (23), and (ii) the sample mean and standard deviation of the real stock price, the real and return volatility. The sample covers monthly data for the period from January 1950 to December 2006.

	Estimate	Std error
g	0.0381	0.0171
σ_G^2	0.0012	$5.2089 \cdot 10^{-4}$
s_0	0.2272	0.1154
s_1	−0.8956	0.4426
s_2	1.8925	0.8873
s_3	−0.0560	0.0292
κ_3	0.0101	0.0043
μ_3	1	restricted
α_3	1.2009	0.4699
β_3	0.0098	0.0106

Table 3

Parameter estimates and block-bootstrap standard errors for the risk-premium parameters of the total risk-premium process in Eq. (7):

$$\begin{aligned}
\pi_1(y_1(t)) &= \alpha_1\lambda_{1(1)} + \left(\beta_1\lambda_{1(1)} + \lambda_{2(1)}\right) y_1(t) && \text{(inflation)} \\
\pi_2(y_2(t)) &= \alpha_2\lambda_{1(2)} + \left(\beta_2\lambda_{1(2)} + \lambda_{2(2)}\right) y_2(t) && \text{(industrial production)} \\
\pi_3(y_3(t)) &= \alpha_3\lambda_{1(3)} + \left(\beta_3\lambda_{1(3)} + \lambda_{2(3)}\right) y_3(t) && \text{(unobservable factor)}
\end{aligned}$$

where $y_1(t)$ and $y_2(t)$ are gross inflation and gross industrial production growth, as defined in Table 1, and $y_3(t)$ is the unobserved factor. The parameter vector is $\boldsymbol{\lambda}^\top = \left(\lambda_{1(1)}, \lambda_{2(1)}, \lambda_{1(2)}, \lambda_{2(2)}, \lambda_{1(3)}, \lambda_{2(3)}\right)$. Parameter estimates are obtained through the third step of the estimation procedure set forth in Section 3.3, relying on Indirect Inference and Simulated Method of Moments. Matching conditions relate to (i) parameter estimates for the auxiliary model for the VIX index, Eq. (25), and (ii) the sample mean and standard deviation of the VIX index. The sample covers monthly data for the period from January 1990 to December 2006.

	Estimate	Std error
Inflation		
$\lambda_{1(1)}$	$-2.1533\cdot10^3$	$0.9683\cdot10^3$
$\lambda_{2(1)}$	32.0141	15.5655
Ind. Prod.		
$\lambda_{1(2)}$	$5.6760\cdot10^2$	$2.7643\cdot10^2$
$\lambda_{2(2)}$	5.5717	2.7952
Unobs.		
$\lambda_{1(3)}$	0.0019	0.0008
$\lambda_{2(3)}$	$5.9837\cdot10^{-4}$	$2.9526\cdot10^{-4}$

Table 4

Variance decomposition statistics for (i) realized volatility (Panel A) and expected volatility under the risk-neutral probability (Panel B). Panel A reports averages and standard deviations of the contributions to the total variance, $\sigma^2(t)$ in Eq. (12), made by: (i) the two macroeconomic factors, gross inflation, $y_1(t)$, and gross industrial production growth, $y_2(t)$, as defined in Table 1, (ii) the unobserved factor, $y_3(t)$, and (iii) secular growth, defined respectively, as:

$$C_j(t) \equiv \frac{s_j^2\left(\alpha_j + \beta_j y_j(t)\right)}{\sigma^2(t)}, \; j = 1, 2, 3, \quad \text{and} \quad C_G(t) \equiv \frac{\sigma_G^2}{\sigma^2(t)}.$$

Paths for the contributions $C_j(t)$ and $C_G(t)$ are generated by feeding them with the two macroeconomic factors $y_1(t)$ and $y_2(t)$ and by averaging over the cross-section of 1000 dynamic simulations of the unobserved factor. The sample covers monthly data for the period from January 1950 to December 2006. Panel B reports statistics for the risk-neutral counterparts to the average paths of $C_j(t)$ and $C_G(t)$. The sample covers monthly data for the period from January 1990 to December 2006.

Panel A: Contributions of factors to stock volatility

Averages	1950-2006	1950-1981	1960-1981	1982-2006
Gross inflation	0.87%	0.92%	0.88%	0.83%
Gross growth	73.47%	71.97%	73.57%	75.10%
Unobserved factor	17.23%	18.10%	17.41%	16.27%
Secular growth	8.43%	9.01%	8.13%	7.78%

Standard deviations	1950-2006	1950-1981	1960-1981	1982-2006
Gross inflation	0.18%	0.23%	0.12%	0.08%
Gross growth	7.53%	9.65%	5.03%	3.47%
Unobserved factor	3.69%	4.67%	2.46%	1.71%
Secular growth	3.67%	4.75%	2.40%	1.68%

Panel B: Contributions of factors to the VIX Index

	Averages	Standard deviations
Gross inflation	1.55%	0.03%
Gross growth	86.91%	1.28%
Unobserved factor	8.79%	0.56%
Secular growth	2.75%	0.17%

Figures

1

1

1

1

1

Supplemental material: technical appendix [Not for publication]

A. Supplemental material for Section 2

A multifactor model

The model we consider differs from those in Bekaert and Grenadier (2001), Ang and Liu (2004) or Mamaysky (2002), for a number of reasons. First, we consider a continuous-time framework, which avoids theoretical challenges pointed out by Bekaert and Grenadier (2001). Furthermore, Ang and Liu (2004) consider a discrete-time setting in which expected returns are exogenous, while in our model, expected returns are endogenous. Finally, our model works differently from Mamaysky's because it endogenously determines the price-dividend ratio.

We consider a multifactor model where a vector-valued process $\boldsymbol{y}(t)$ is solution to a n-dimensional affine diffusion,

$$\mathrm{d}\boldsymbol{y}(t) = \boldsymbol{\kappa}(\boldsymbol{\mu} - \boldsymbol{y}(t))\,\mathrm{d}t + \boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y}(t))\,\mathrm{d}\boldsymbol{W}(t), \tag{29}$$

where $\boldsymbol{W}(t)$ is a d-dimensional Brownian motion ($n \leq d$), $\boldsymbol{\Sigma}$ is a full rank $n \times d$ matrix, and $\boldsymbol{V}$ is a full rank $d \times d$ diagonal matrix with elements,

$$V(\boldsymbol{y})_{(ii)} = \sqrt{\alpha_i + \boldsymbol{\beta}_i^\top \boldsymbol{y}}, \quad i = 1, \cdots, d,$$

for some scalars α_i and vectors $\boldsymbol{\beta}_i$. We assume that the Brownian motion driving secular growth, $W_G(t)$ in Eq. (4), is uncorrelated with $\boldsymbol{W}(t)$ in Eq. (29). We shall review soon sufficient conditions known to ensure that Eq. (29) has a strong solution with $\boldsymbol{V}(\boldsymbol{y}(t))_{(ii)} > 0$ almost surely for all t.

The model we estimate, Eq. (1) in Section 2 of the main text, is a special case of Eq. (29), with $n = d = 3$, the matrix $\boldsymbol{\kappa}$ given by:

$$\boldsymbol{\kappa} = \begin{bmatrix} \kappa_1 & \bar{\kappa}_1 & 0 \\ \bar{\kappa}_2 & \kappa_2 & 0 \\ 0 & 0 & \kappa_3 \end{bmatrix},$$

and with $\boldsymbol{\Sigma} = \boldsymbol{I}_{3\times 3}$ and the vectors $\boldsymbol{\beta}_i$ being such that $\beta_j \equiv \beta_{jj}$.

While one does not necessarily observe every single component of $\boldsymbol{y}(t)$, we do observe discretely sampled paths of macroeconomic variables such as industrial production, unemployment or inflation. Let $\{\mathrm{M}_{j,t}\}_{t=1,2,\cdots}$ be the discretely sampled path of the macroeconomic variable $\mathrm{M}_{j,t}$ where, for example, $\mathrm{M}_{j,t}$ can be the industrial production index available for month t, and $j = 1, \cdots, N_{\mathrm{M}}$, where N_{M} is the number of observed macroeconomic factors. We assume, without loss of generality, that these observed macroeconomic factors are strictly positive, and that they are related to the state vector process in Eq. (29) by:

$$\ln(\mathrm{M}_{j,t}/\,\mathrm{M}_{j,t-12}) = f_j(\boldsymbol{y}(t)), \quad j = 1, \cdots, N_{\mathrm{M}}, \tag{30}$$

where the collection of functions $\{f_j\}$ determines how the factors dynamics impinge upon the observed macroeconomic variables. In terms of the model in the main text, the functions in Eq. (30) are $f_j(\boldsymbol{y}) \equiv \ln y_j$.

We now turn to model asset prices. We assume that asset prices are related to the vector of factors $\boldsymbol{y}(t)$ in Eq. (29), and that some of these factors affect developments in macroeconomic conditions, through Eq. (30). For analytical convenience, we rule out that asset prices can feed back the real economy, although we acknowledge that the presence of frictions can make capital markets and the macroeconomy intimately related, as in the financial accelerator hypothesis reviewed by Bernanke, Gertler and Gilchrist (1999), or in the static model analyzed by Angeletos, Lorenzoni and Pavan (2008), where feedbacks arise due to asymmetric information and learning between agents acting within the real and the financial spheres of the economy.

The Arrow-Debreu density we consider is exactly that in Eq. (5), with the sole exception that the vector Brownian motion $\boldsymbol{W}$ is the d-dimensional one in Eq. (29). Consider, then, the following "essentially affine" specification for the dynamics of the factors in Eq. (29), and the risk-premiums. Let $\boldsymbol{V}^-(\boldsymbol{y})$ be a $d \times d$ diagonal matrix with elements

$$V^-(\boldsymbol{y})_{(ii)} = \begin{cases} \frac{1}{V(\boldsymbol{y})_{(ii)}} & \text{if } \Pr\{V(\boldsymbol{y}(t))_{(ii)} > 0 \text{ all } t\} = 1 \\ 0 & \text{otherwise} \end{cases}$$

and set, $\boldsymbol{\Lambda}(\boldsymbol{y}) = \boldsymbol{V}(\boldsymbol{y})\boldsymbol{\lambda}_1 + \boldsymbol{V}^-(\boldsymbol{y})\boldsymbol{\lambda}_2\boldsymbol{y}$, for some d-dimensional vector $\boldsymbol{\lambda}_1$ and some $d \times n$ matrix $\boldsymbol{\lambda}_2$.

By the definition of the dividends in Eq. (2), the stock price follows:

$$\frac{\mathrm{d}S(t)}{S(t)}=\left(r-\frac{G(t)\,\delta(\boldsymbol{y}(t))}{S(\boldsymbol{y}(t))}\right)\mathrm{d}t+\frac{s_y(\boldsymbol{y}(t))^\top\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))}\mathrm{d}\hat{\boldsymbol{W}}(t)+\sigma_G\mathrm{d}\hat{W}_G(t), \tag{31}$$

where $\hat{\boldsymbol{W}}$ and $\hat{W}_G$ are Brownian motions defined under the risk-neutral probability Q. Under regularity conditions provided below, and in the absence of bubbles, Eq. (31) implies that the stock price is,

$$S(G,\boldsymbol{y})=\mathbb{E}\left[\int_t^\infty e^{-r(s-t)}G(s)\,\delta(\boldsymbol{y}(s))\,\mathrm{d}s\,\middle|\,G(t)=G,\boldsymbol{y}(t)=\boldsymbol{y}\right], \tag{32}$$

where $\mathbb{E}$ is the expectation taken under the risk-neutral probability Q.

We are only left with specifying how the instantaneous dividend relates to the state vector $\boldsymbol{y}$. Let

$$\delta(\boldsymbol{y})=\delta_0+\boldsymbol{\delta}^\top\boldsymbol{y}, \tag{33}$$

for some scalar δ_0 and some vector $\boldsymbol{\delta}$.

We have:

Proposition A1: *Let the risk-premiums be as in Eq. (6), and the instantaneous dividend rate be as in Eqs. (2) and (33). Then, under a technical regularity condition (condition (38)), we have that: (i) Eq. (32) holds; and (ii) the rational stock price function $S(G,\boldsymbol{y})=G\cdot s(\boldsymbol{y})$, where $s(\boldsymbol{y})$ is affine in the state vector $\boldsymbol{y}$, viz*

$$s(\boldsymbol{y})=\frac{\delta_0+\boldsymbol{\delta}^\top(\boldsymbol{D}+(r-g+\sigma_G\lambda_G)\boldsymbol{I}_{n\times n})^{-1}\boldsymbol{c}}{r-g+\sigma_G\lambda_G}+\boldsymbol{\delta}^\top(\boldsymbol{D}+(r-g+\sigma_G\lambda_G)\boldsymbol{I}_{n\times n})^{-1}\boldsymbol{y}, \tag{34}$$

where

$$\boldsymbol{c}=\boldsymbol{\kappa\mu}-\boldsymbol{\Sigma}\begin{pmatrix}\alpha_1\lambda_{1(1)} & \cdots & \alpha_d\lambda_{1(d)}\end{pmatrix}^\top \tag{35}$$

$$\boldsymbol{D}=\boldsymbol{\kappa}+\boldsymbol{\Sigma}\left[\begin{pmatrix}\lambda_{1(1)}\boldsymbol{\beta}_1^\top & \cdots & \lambda_{1(d)}\boldsymbol{\beta}_d^\top\end{pmatrix}^\top+\boldsymbol{I}^-\boldsymbol{\lambda}_2\right], \tag{36}$$

$\boldsymbol{I}^-$ is a $d\times d$ diagonal matrix with elements $I^-_{(ii)}=1$ if $\Pr\{V(\boldsymbol{y}(t))_{(ii)}>0$ all $t\}=1$ and 0 otherwise; and, finally $\{\lambda_{1(j)}\}_{j=1}^d$ are the components of $\boldsymbol{\lambda}_1$.

Existence of a strong solution to Eq. (29)

Consider the following conditions: for all i,

(i) For all $\boldsymbol{y}: V(\boldsymbol{y})_{(ii)}=0$, $\boldsymbol{\beta}_i^\top(-\boldsymbol{\kappa y}+\boldsymbol{\kappa\mu})>\frac{1}{2}\boldsymbol{\beta}_i^\top\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top\boldsymbol{\beta}_i$

(ii) For all j, if $\left(\boldsymbol{\beta}_i^\top\boldsymbol{\Sigma}\right)_j\neq 0$, then $V_{ii}=V_{jj}$.

Then, by Duffie and Kan (1996) (unnumbered theorem, p. 388), there exists a unique strong solution to Eq. (29) for which $V(\boldsymbol{y}(t))_{(ii)}>0$ for all t almost surely.

We apply these conditions to the diffusion in Eq. (1). Condition (i) collapses to,

$$\text{For all } y_i:\alpha_i+\beta_i y_i=0,\quad \beta_i\left[\kappa_i(\mu_i-y_i)+\bar{\kappa}_i(\mu_j-y_j)\right]>\frac{1}{2}\beta_i^2,\quad i\neq j,$$

with $\bar{\kappa}_3\equiv 0$. That is, ruling out the trivial case $\beta_i=0$,

$$\kappa_i(\mu_i\beta_i+\alpha_i)+\bar{\kappa}_i\beta_i\left(\mu_j+\frac{\alpha_j}{\beta_j}\right)>\frac{1}{2}\beta_i^2,\quad i\neq j. \tag{37}$$

Proof of Proposition A1

The technical condition in Proposition A1 is,

$$E\left[\int_t^T \left\| \frac{\boldsymbol{\eta}^\top \boldsymbol{\Sigma} \boldsymbol{V}(\boldsymbol{y}(\tau))}{\gamma + \boldsymbol{\eta}^\top \boldsymbol{y}(\tau)} - \boldsymbol{\Lambda}(\tau)^\top \right\|^2 \mathrm{d}\tau\right] < \infty, \tag{38}$$

for some constants γ and $\boldsymbol{\eta}$ in Eq. (47) below.

We proceed as follows. First, we determine the solution to the stock price, in the absence of secular growth, i.e. when

$$g = \sigma_G \equiv 0. \tag{39}$$

Then, we generalize, by elaborating on Eq. (32) of the main text, as in Eq. (48) below.

When Eq. (39) holds true, define the Arrow-Debreu adjusted asset price process as, $s^\xi(t) \equiv e^{-rt}\xi(t)\, s(\boldsymbol{y}(t))$, $t > 0$. By Itô's lemma, it satisfies,

$$\frac{\mathrm{d}s^\xi(t)}{s^\xi(t)} = \mathrm{Dr}(\boldsymbol{y}(t))\,\mathrm{d}t + \left(\boldsymbol{Q}(\boldsymbol{y}(t))^\top - \boldsymbol{\Lambda}(\boldsymbol{y}(t))^\top\right)\mathrm{d}\boldsymbol{W}(t), \tag{40}$$

where

$$\begin{aligned}
\mathrm{Dr}(\boldsymbol{y}) &= -r + \frac{\mathcal{A}s(\boldsymbol{y})}{s(\boldsymbol{y})} - \boldsymbol{Q}(\boldsymbol{y})^\top \boldsymbol{\Lambda}(\boldsymbol{y}), \\
\mathcal{A}s(\boldsymbol{y}) &= s_y(\boldsymbol{y})^\top \boldsymbol{\kappa}(\boldsymbol{\mu} - \boldsymbol{y}) + \frac{1}{2}\mathrm{Tr}\left([\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})][\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})]^\top s_{yy}(\boldsymbol{y})\right), \quad \boldsymbol{Q}(\boldsymbol{y})^\top = \frac{s_y(\boldsymbol{y})^\top \boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})}{s(\boldsymbol{y})},
\end{aligned}$$

and s_y and s_{yy} denote the gradient and the Hessian of s with respect to $\boldsymbol{y}$. By absence of arbitrage opportunities, for any $T < \infty$,

$$s^\xi(t) = E\left[\int_t^T \delta^\xi(h)\,\mathrm{d}h \,\middle|\, \mathbb{F}(t)\right] + E[s^\xi(T) \mid \mathbb{F}(t)], \tag{41}$$

where $\delta^\xi(t)$ is the current Arrow-Debreu value of the dividend to be paid off at time t, viz $\delta^\xi(t) = e^{-rt}\xi(t)\,\delta(t)$. Below, we show that the following transversality condition holds,

$$\lim_{T\to\infty} E[s^\xi(T) \mid \mathbb{F}(t)] = 0, \tag{42}$$

from which Eq. (32) in the main text follows, once we show that $\int_t^\infty E[\delta^\xi(h)]\mathrm{d}h < \infty$.

Next, by Eq. (41),

$$0 = \frac{\mathrm{d}}{\mathrm{d}\tau} E[s^\xi(\tau) \mid \mathbb{F}(t)]\Big|_{\tau=t} + \delta^\xi(t). \tag{43}$$

Below, we show that

$$E[s^\xi(T) \mid \mathbb{F}(t)] = s^\xi(t) + \int_t^T \mathrm{D}(\boldsymbol{y}(h))\, s^\xi(h)\,\mathrm{d}h. \tag{44}$$

Therefore, by the assumptions on $\boldsymbol{\Lambda}$, Eq. (43) can be rearranged to yield the following ordinary differential equation,

$$\text{For all } \boldsymbol{y}, \quad s_y(\boldsymbol{y})^\top(\boldsymbol{c} - \boldsymbol{D}\boldsymbol{y}) + \frac{1}{2}\mathrm{Tr}\left([\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})][\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})]^\top s_{yy}(\boldsymbol{y})\right) + \delta(\boldsymbol{y}) - rs(\boldsymbol{y}) = 0, \tag{45}$$

where $\boldsymbol{c}$ and $\boldsymbol{D}$ are defined in the proposition.

Assume that the price function is affine in $\boldsymbol{y}$,

$$s(\boldsymbol{y}) = \gamma + \boldsymbol{\eta}^\top \boldsymbol{y}, \tag{46}$$

for some scalar γ and some vector $\boldsymbol{\eta}$. By plugging this guess back into Eq. (45) we obtain,

$$\text{For all } \boldsymbol{y}, \quad \boldsymbol{\eta}^\top \boldsymbol{c} + \delta_0 - r\gamma - \left[\boldsymbol{\eta}^\top(\boldsymbol{D} + r\boldsymbol{I}_{n\times n}) - \boldsymbol{\delta}^\top\right]\boldsymbol{y} = 0.$$

That is,

$$\boldsymbol{\eta}^\top \boldsymbol{c} + \delta_0 - r\gamma = 0 \quad \text{and} \quad \left[\boldsymbol{\eta}^\top \left(\boldsymbol{D} + r\boldsymbol{I}_{n\times n}\right) - \boldsymbol{\delta}^\top\right] = \boldsymbol{0}_{1\times n}.$$

The solution to this system is,

$$\gamma = \frac{\delta_0 + \boldsymbol{\eta}^\top \boldsymbol{c}}{r} \quad \text{and} \quad \boldsymbol{\eta}^\top = \boldsymbol{\delta}^\top \left(\boldsymbol{D} + r\boldsymbol{I}_{n\times n}\right)^{-1}. \tag{47}$$

We are left to show that Eq. (42) and (44) hold true, when Eq. (39) also holds true.

As for Eq. (42), we have:

$$\begin{aligned}\lim_{T\to\infty} E[s^\xi(T) \mid \mathbb{F}(t)] &= \lim_{T\to\infty} E[e^{-r(T-t)}\xi(T)\, s(\boldsymbol{y}(T)) \mid \mathbb{F}(t)] \\ &= \gamma \lim_{T\to\infty} e^{-r(T-t)} E[\xi(T) \mid \mathbb{F}(t)] + \lim_{T\to\infty} e^{-r(T-t)} E[\xi(T)\, \boldsymbol{\eta}^\top \boldsymbol{y}(T) \mid \mathbb{F}(t)] \\ &= \xi(t) \lim_{T\to\infty} e^{-r(T-t)} \mathbb{E}[\boldsymbol{\eta}^\top \boldsymbol{y}(T) \mid \mathbb{F}(t)],\end{aligned}$$

where the second line follows by Eq. (46), and the third line holds because $E[\xi(T) \mid \mathbb{F}(t)] = 1$ for all T, and by a change of measure. Eq. (42) follows because $\boldsymbol{y}$ is stationary mean-reverting under the risk-neutral probability.

To show that Eq. (44) holds, we need to show that the diffusion part of s^ξ in Eq. (40) is a martingale, not only a local martingale, which it does whenever for all T,

$$E\left[\int_t^T \left\|\boldsymbol{Q}(\boldsymbol{y}(\tau))^\top - \boldsymbol{\Lambda}(\tau)^\top\right\|^2 \mathrm{d}\tau\right] < \infty,$$

which is the condition in (38). This ends the proof of Proposition A1, in the case $g = \sigma_G \equiv 0$.

For the general case of Proposition A1, note that by Eq. (32):

$$\begin{aligned}&S(G, \boldsymbol{y}) \\ &= G \cdot \mathbb{E}\left[\int_t^\infty e^{-r(s-t)} \mathbb{E}\left(\frac{G(s)}{G}\delta(\boldsymbol{y}(s))\,\mathrm{d}s \,\middle|\, G(t) = G\right) \middle|\, G(t) = G, \boldsymbol{y}(t) = \boldsymbol{y}\right] \\ &= G \cdot \mathbb{E}\left[\int_t^\infty e^{-r(s-t)} \mathbb{E}\left(e^{\left(g - \frac{1}{2}\sigma_G^2 - \lambda_G\sigma_G\right)(s-t) + \sigma_G\left(\tilde{W}_G(s) - \tilde{W}_G(t)\right)} \,\middle|\, G(t) = G\right)\delta(\boldsymbol{y}(s))\,\mathrm{d}s \,\middle|\, \boldsymbol{y}(t) = \boldsymbol{y}\right] \\ &= G \cdot \mathbb{E}\left[\int_t^\infty e^{-(r - g + \lambda_G\sigma_G)(s-t)}\delta(\boldsymbol{y}(s))\,\mathrm{d}s \,\middle|\, \boldsymbol{y}(t) = \boldsymbol{y}\right],\end{aligned} \tag{48}$$

where the first equality follows by the law of iterated expectations, the second by the independence of G and $\boldsymbol{y}$, and the definition of G in Eq. (4) of the main text, and the third from a simple computation. The term in the brackets is the same as the RHS of Eq. (32) of the main text, for $G(s) \equiv 1$, $s \in (t, \infty)$. Therefore, the solution for the term in the brackets is the same as that provided in the case of absence of secular growth, i.e. when Eq. (39) holds true, but with $r - g + \lambda_G\sigma_G$ replacing r.

B. Supplemental material for Section 3

Remarks on notation: Hereafter, we let Avar and Acov denote the limits of the variance and covariance operators, respectively. Let $\boldsymbol{u}$ be a $n \times 1$ vector, where each element depends on some $m \times 1$ parameter vector $\boldsymbol{\theta}$. We define: the $m \times n$ matrix $\nabla_\theta \boldsymbol{u} = \frac{\partial \boldsymbol{u}^\top}{\partial \boldsymbol{\theta}}$; $\|\boldsymbol{u}\|^p = \left(\sqrt{\boldsymbol{u}^\top \boldsymbol{u}}\right)^p$, for some scalar $p > 0$; and $|\boldsymbol{u}|_2 = \boldsymbol{u}\boldsymbol{u}^\top$, the outer product of $\boldsymbol{u}$. Finally, for any $n \times m$ matrix $\boldsymbol{A}$, we set $|\boldsymbol{A}| = \sum_{i=1}^n \sum_{j=1}^m |a_{i,j}|$.

B.1. Asymptotic theory for the estimators in Section 3

The sets $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$ in Sections 3.1 and 3.2 are defined as:

$$\boldsymbol{\Phi} = \left\{\boldsymbol{\phi} : \text{The inequality in (37) holds, } \kappa_i > 0, \text{ and } \kappa_i\kappa_j - \bar{\kappa}_i\bar{\kappa}_j > 0,\ i,j = 1,2 \text{ and } i \neq j\right\},$$

and

$$\boldsymbol{\Theta} = \left\{\boldsymbol{\theta} : \text{The inequality in (37) holds for } i = 3, \text{ and } \kappa_3 > 0\right\}.$$

Furthermore, we let $\boldsymbol{\phi}_0$ and $\boldsymbol{\theta}_0$ be the solutions to the two limit problems,

$$\boldsymbol{\phi}_0 = \arg\min_{\boldsymbol{\phi}\in\boldsymbol{\Phi}_0} \operatorname*{plim}_{T\to\infty,\Delta\to 0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}\left(\boldsymbol{\phi}\right) - \tilde{\boldsymbol{\varphi}}_T \right\|^2, \tag{49}$$

and

$$\boldsymbol{\theta}_0 = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \operatorname*{plim}_{T\to\infty,\Delta\to 0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}\right) - \tilde{\boldsymbol{\vartheta}}_T \right\|^2, \tag{50}$$

where $\boldsymbol{\Phi}_0$ and $\boldsymbol{\Theta}_0$ are compact sets of $\boldsymbol{\Phi}$ and $\boldsymbol{\Theta}$, respectively. Finally, we define the limit problem for the estimator of the risk-premium parameters,

$$\boldsymbol{\lambda}_0 = \arg\min_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}_0} \operatorname*{plim}_{T\to\infty,\Delta\to 0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \boldsymbol{\lambda}) - \tilde{\boldsymbol{\psi}}_T \right\|^2. \tag{51}$$

We are now ready to analyze the asymptotic behavior of these estimators. The following assumption summarizes the properties of the data generating mechanism we rely on.

Assumption B1: (i) Conditions (i) and (ii) in Appendix A hold for $i = 1, 2, 3$; (ii) The sample observations for the macroeconomic factors $y_1(t), y_2(t)$ are generated by Eq. (1) for $j = 1, 2$; (iii) As for Eq. (1), for $i, j = 1, 2$ $i \neq j$, $\kappa_i\kappa_j - \bar{\kappa}_i\bar{\kappa}_j > 0$ and for all $i = 1, 2, 3$ $\kappa_i > 0$; (iv) The sample observations for the stock market index $s(t)$ are generated by Eq. (8); (v) The risk-premium vector $\boldsymbol{\pi}(\boldsymbol{y})$ and the dividend vector $\delta(\boldsymbol{y})$ are defined as in Eqs. (7) and (3).

The estimator of $\hat{\boldsymbol{\phi}}_T$ in Eq. (19)

We have:

Proposition B1: *Under regularity conditions (Assumption B1(i)-(iii) in Appendix B), as $T \to \infty$ and $\Delta\sqrt{T} \to 0$,*

$$\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0\right) \xrightarrow{d} \mathrm{N}\left(\boldsymbol{0}, \boldsymbol{V}_1\right), \quad \boldsymbol{V}_1 = \left(1 + \frac{1}{H}\right)\left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1} \boldsymbol{D}_1^\top \boldsymbol{J}_1 \boldsymbol{D}_1 \left(\boldsymbol{D}_1^\top \boldsymbol{D}_1\right)^{-1},$$

where $\boldsymbol{\phi}_0$ is as in Eq. (49), and the two matrices, $\boldsymbol{D}_1$ and $\boldsymbol{J}_1$, are defined in the proof below.

Proof: By the conditions in Assumptions B1(i) and B1(ii), $(y_1(t), y_2(t))$ admits a unique strong solution, and has a positive-definite covariance matrix with probability one. Assumption B1(iii) ensures that $(y_1(t), y_2(t))$ is geometrically ergodic and the skeleton $(y_{1,t}, y_{2,t})$ is geometrically β-mixing. Further, by Glasserman and Kim (2010), the stationary distribution of $(y_1(t), y_2(t))$ and $(y_{1,t}, y_{2,t})$ has exponential tails, which ensures that there are enough finite moments for the uniform law of large numbers and the central limit theorem to apply. By the same argument, for any $\boldsymbol{\phi} \in \boldsymbol{\Phi}_0$, the simulated skeleton $(y^{\phi}_{1,t,\Delta,h}, y^{\phi}_{2,t,\Delta,h})$ is also geometrically β-mixing, with stationary distribution having exponential tails. Finally, given Eq. (1), $(y^{\phi}_{1,t,\Delta,h}, y^{\phi}_{2,t,\Delta,h})$ is at least twice continuously differentiable in any open neighborhood of $\boldsymbol{\phi}_0$.

We have that $\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0 = o_p(1)$, because of the uniform law of large numbers and unique identifiability. Next, by the first order conditions and a mean-value expansion around $\boldsymbol{\phi}_0$,

$$\begin{aligned} 0 &= \nabla_\phi \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T)\right)^\top \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T) - \tilde{\boldsymbol{\varphi}}_T\right) \\ &= \nabla_\phi \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T)\right)^\top \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}\left(\boldsymbol{\phi}_0\right) - \tilde{\boldsymbol{\varphi}}_T\right) \\ &+ \nabla_\phi \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T)\right)^\top \nabla_\phi \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\bar{\boldsymbol{\phi}}_T)\right)\left(\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0\right), \end{aligned}$$

where $\bar{\boldsymbol{\phi}}_T$ is some convex combination of $\hat{\boldsymbol{\phi}}_T$ and $\boldsymbol{\phi}_0$. Let

$$\boldsymbol{D}_1\left(\boldsymbol{\phi}_0\right) \equiv \boldsymbol{D}_1 = \operatorname{plim} \nabla_\phi \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}\left(\boldsymbol{\phi}_0\right)\right).$$

By the uniform law of large numbers, $\sup_{\boldsymbol{\phi}\in\boldsymbol{\Phi}_0}\left|\nabla_\phi\left(\frac{1}{H}\sum_{h=1}^H\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi})\right)-\boldsymbol{D}_1(\boldsymbol{\phi})\right|=o_p(1)$, and as $\hat{\boldsymbol{\phi}}_T-\boldsymbol{\phi}_0=o_p(1)$, $\nabla_\phi\left(\frac{1}{H}\sum_{h=1}^H\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\bar{\boldsymbol{\phi}}_T)\right)-\boldsymbol{D}_1=o_p(1)$. Hence,

$$\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T-\boldsymbol{\phi}_0\right)=-\left(\boldsymbol{D}_1^\top\boldsymbol{D}_1\right)^{-1}\boldsymbol{D}_1^\top\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)-\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T-\boldsymbol{\varphi}_0\right)\right)+o_p(1).$$

Let $\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)$ be the unfeasible estimator, obtained by simulating continuous paths for $y_j(t)$, i.e. $y_{j,t,h}^{\phi_0}$, $j=1,2$. We claim that for $h=1,\cdots,H$,

$$\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi}_0)-\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)\right)=o_p(1).$$

Let $\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}$ be the vector containing all the regressors in Eq. (18), and let $\hat{\boldsymbol{\varphi}}_{1,T,\Delta,h}(\boldsymbol{\phi}_0)$ be the parameter estimator of the OLS regression of $y_{1,t,\Delta,h}^{\phi_0}$ on $\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}$. We have:

$$\begin{aligned}
&\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{1,T,\Delta,h}(\boldsymbol{\phi}_0)-\hat{\boldsymbol{\varphi}}_{1,T,h}(\boldsymbol{\phi}_0)\right)\\
&=\left(\frac{1}{T}\sum_{t=25}^T\boldsymbol{Y}_{t,h}^{\phi_0}\boldsymbol{Y}_{t,h}^{\phi_0\top}\right)^{-1}\sqrt{T}\left(\frac{1}{T}\sum_{t=25}^T\left(\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}y_{1,t,\Delta,h}^{\phi_0}-\boldsymbol{Y}_{t,h}^{\phi_0}y_{1,t,h}^{\phi_0}\right)\right)\\
&+\sqrt{T}\left(\left(\frac{1}{T}\sum_{t=25}^T\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}\boldsymbol{Y}_{t,\Delta,h}^{\phi_0\top}\right)^{-1}-\left(\frac{1}{T}\sum_{t=25}^T\boldsymbol{Y}_{t,h}^{\phi_0}\boldsymbol{Y}_{t,h}^{\phi_0\top}\right)^{-1}\right)\left(\frac{1}{T}\sum_{t=25}^T\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}y_{1,t,\Delta,h}^{\phi_0}\right).
\end{aligned}\tag{52}$$

As for the first term on the RHS of (52), $\left(\frac{1}{T}\sum_{t=25}^T\boldsymbol{Y}_{t,h}^{\phi_0}\boldsymbol{Y}_{t,h}^{\phi_0\top}\right)^{-1}=O_p(1)$, and by Theorem 2.3 in Pardoux and Talay (1985), we have, for $\varepsilon>0$ and $\sqrt{T}\Delta\to 0$,

$$\Pr\left(\left|\frac{1}{\sqrt{T}}\sum_{t=25}^T\left(\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}y_{1,t,\Delta,h}^{\phi_0}-\boldsymbol{Y}_{t,h}^{\phi_0}y_{1,t,h}^{\phi_0}\right)\right|>\varepsilon\right)<\frac{1}{\varepsilon}\sqrt{T}\mathrm{E}\left(\left|\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}y_{1,t,\Delta,h}^{\phi_0}-\boldsymbol{Y}_{t,h}^{\phi_0}y_{1,t,h}^{\phi_0}\right|\right)=\sqrt{T}O(\Delta)=o(1).$$

The second term on the RHS of Eq. (52) can be dealt with similarly. Thus, we have:

$$\mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T-\boldsymbol{\phi}_0\right)\right)=\left(\boldsymbol{D}_1^\top\boldsymbol{D}_1\right)^{-1}\boldsymbol{D}_1^\top\mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)-\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T-\boldsymbol{\varphi}_0\right)\right)\boldsymbol{D}_1\left(\boldsymbol{D}_1^\top\boldsymbol{D}_1\right)^{-1},$$

where,

$$\begin{aligned}
&\mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)-\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T-\boldsymbol{\varphi}_0\right)\right)\\
&=\mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)\right)+\mathrm{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T-\boldsymbol{\varphi}_0\right)\right)\\
&-2\mathrm{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right),\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T-\boldsymbol{\varphi}_0\right)\right).
\end{aligned}$$

The last term of the RHS of this equality is zero, because the simulated paths are independent of the sample paths. Moreover, the simulated paths are independent and identically distributed across all simulation replications and, hence,

$$\mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)\right)=\frac{1}{H}\mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)\right),\text{ for all } h.$$

Finally, given Assumption B1(ii),

$$\boldsymbol{J}_1\equiv\mathrm{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T-\boldsymbol{\varphi}_0\right)\right)=\mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)\right),\text{ for all } h,$$

and so

$$\mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T-\boldsymbol{\phi}_0\right)\right)=\left(1+\frac{1}{H}\right)\left(\boldsymbol{D}_1^\top\boldsymbol{D}_1\right)^{-1}\boldsymbol{D}_1^\top\boldsymbol{J}_1\boldsymbol{D}_1\left(\boldsymbol{D}_1^\top\boldsymbol{D}_1\right)^{-1}.$$

1 The proposition follows by the central limit theorem for geometrically strong mixing processes.

The estimator of $\hat{\boldsymbol{\theta}}_T$ in Eq. (24)

We have:

Proposition B2: *Under regularity conditions (Assumption B1(i)-(iv) in Appendix B), as $T \to \infty$ and $\Delta\sqrt{T} \to 0$,*

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right) \xrightarrow{d} \mathrm{N}\left(\mathbf{0}, \boldsymbol{V}_2\right), \quad \boldsymbol{V}_2 = \left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} \boldsymbol{D}_2^\top \left(\left(1 + \frac{1}{H}\right)\left(\boldsymbol{J}_2 - \boldsymbol{K}_2\right) + \boldsymbol{P}_2\right) \boldsymbol{D}_2 \left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1},$$

where $\boldsymbol{\theta}_0$ is as in Eq. (50), and the four matrices, $\boldsymbol{D}_2$, $\boldsymbol{J}_2$, $\boldsymbol{K}_2$ and $\boldsymbol{P}_2$, are defined in the proof below.

As discussed in the main text, the matrix $\boldsymbol{P}_2$ arises due to parameter estimation error, as the stock price in Eq. (20), is simulated with parameters $\boldsymbol{\theta}$ fixed at their estimates, $\hat{\boldsymbol{\theta}}_{G,T}$. Moreover, the matrix $\boldsymbol{K}_2$ captures the covariance of the structural parameter estimates over all the simulation replications, as well as the covariance between actual and simulated paths, thereby resulting in an improved efficiency, if compared to estimators based on unconditional (simulated) inference.

Proof of Proposition B2: By the same arguments utilized in the proof of Proposition B1,

$$\begin{aligned}\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right) = &-\left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} \boldsymbol{D}_2^\top \left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right) + \boldsymbol{C}_2^\top \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T} - \boldsymbol{\theta}_G\right)\right.\\ &\left.-\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right) + o_p(1),\end{aligned}$$

where for $\bar{\boldsymbol{\theta}}_{G,T} \in (\hat{\boldsymbol{\theta}}_{G,T}, \boldsymbol{\theta}_G)$,

$$\boldsymbol{D}_2 = \mathrm{plim}\ \nabla_\theta \left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \bar{\boldsymbol{\theta}}_{G,T}\right)\right), \quad \boldsymbol{C}_2 = \mathrm{plim}\ \nabla_{\theta_G} \left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \bar{\boldsymbol{\theta}}_{G,T}\right)\right).$$

Therefore:

$$\mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right)\right) = \left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} \boldsymbol{D}_2^\top \boldsymbol{J}_0 \boldsymbol{D}_2 \left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1},$$

where

$$\begin{aligned}\boldsymbol{J}_0 = &\ \mathrm{Avar}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right) + \mathrm{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right)\\ &+ \boldsymbol{C}_2^\top \mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T} - \boldsymbol{\theta}_G\right)\right)\boldsymbol{C}_2 + 2\boldsymbol{C}_2^\top \mathrm{Acov}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right), \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T} - \boldsymbol{\theta}_G\right)\right)\\ &- 2\mathrm{Acov}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right), \sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right)\\ &- 2\boldsymbol{C}_2^\top \mathrm{Acov}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T} - \boldsymbol{\theta}_G\right), \sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right).\end{aligned} \tag{53}$$

Let $\hat{\boldsymbol{\vartheta}}_{T,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right)$ be the infeasible estimator, obtained by simulating continuous paths for the unobservable factor $y_3\left(t\right)$ and for $G^{\hat{\theta}_{G,T}}(t)$. By the same arguments as those in the proof of Proposition B1,

$$\mathrm{Avar}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right) = \mathrm{Avar}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right).$$

Paths of the model-implied stock price are obtained through the sample paths of the observable factors $y_{1,t}$ and $y_{2,t}$. Therefore, simulated paths are not independent across simulations, and are not independent of the actual sample paths of stock price and volatility. We have:

$$\begin{aligned}&\mathrm{Avar}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right)\\ &= \frac{1}{H}\mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,1}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right)\\ &+ \frac{1}{H^2}\sum_{h=1}^H \sum_{h'=1, h' \neq h}^H \mathrm{Acov}\left(\sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right), \sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,h'}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right)\\ &= \frac{1}{H}\boldsymbol{J}_2 + \frac{H(H-1)}{H^2}\boldsymbol{K}_2,\end{aligned}$$

where

$$\boldsymbol{J}_2 = \text{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right) = \text{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right), \quad \text{for all } h,$$

and

$$\begin{aligned}\boldsymbol{K}_2 &= \frac{1}{H\left(H-1\right)} \sum_{h=1}^{H} \sum_{h'=1, h' \neq h}^{H} \text{Acov}\left(\sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right), \sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,h'}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right) \\ &= \text{Acov}\left(\sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,1}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right), \sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,2}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)^{\top}\right).\end{aligned}$$

Therefore, using the fact that $\text{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right) = \text{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,1}\left(\boldsymbol{\theta}_0\right) - \boldsymbol{\vartheta}_0\right)\right) = \boldsymbol{J}_2$, letting $\boldsymbol{P}_2$ denoting the sum of the third, fourth and sixth terms in Eq. (53), and exploiting the expression for $\boldsymbol{J}_0$, we obtain:

$$\boldsymbol{J}_0 = \frac{1}{H}\boldsymbol{J}_2 + \frac{H(H-1)}{H^2}\boldsymbol{K}_2 + \boldsymbol{J}_2 - 2\boldsymbol{K}_2 + \boldsymbol{P}_2 = \left(1 + \frac{1}{H}\right)\left(\boldsymbol{J}_2 - \boldsymbol{K}_2\right) + \boldsymbol{P}_2,$$

and, hence:

$$\text{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right)\right) = \left(\boldsymbol{D}_2^{\top}\boldsymbol{D}_2\right)^{-1}\boldsymbol{D}_2^{\top}\left(\left(1 + \frac{1}{H}\right)\left(\boldsymbol{J}_2 - \boldsymbol{K}_2\right) + \boldsymbol{P}_2\right)\boldsymbol{D}_2\left(\boldsymbol{D}_2^{\top}\boldsymbol{D}_2\right)^{-1}.$$

Details on the simulations of the VIX index predicted by the model

We construct a simulated series of length $\mathcal{T}$ for the VIX index, at a monthly frequency. Since we do not have a closed-form formula for the VIX index, we need to resort to numerical methods aiming to approximate it. We address this issue by simulating the three factors at a daily frequency, which we then use to numerically integrate the daily volatilities. For each simulation draw $h = 1, \cdots, H$, we initialize each monthly path at the values taken by the observable macroeconomic factors, i.e. at $y_{1,t}$, $y_{2,t}$, $t = T_*, \cdots, T_* + \mathcal{T} - 1$, where T_* is the first date where the VIX is available, and at the monthly unconditional mean of the unobservable factor. For $i = 1, 2, 3$, $h = 1, \cdots, H$, $k = 0, \cdots, \hat{\Delta}^{-1} - 1$, let $\hat{y}^{\lambda}_{i,t+k\hat{\Delta},h}$ be the value of the i-th factor, at time $t + k\hat{\Delta}$, for the h-th simulation under the risk-neutral probability, performed with parameter $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_0$ and remaining parameters fixed at their estimates obtained in the first and second step of our estimation procedure. $\hat{\Delta}$ will be defined in a moment. Simulations are obtained through a Milstein approximation to the risk-neutral version of Eq. (1),

$$\mathrm{d}y_i\left(t\right) = \left[\kappa_i\left(\mu_i - y_i\left(t\right)\right) + \bar{\kappa}_i\left(\bar{\mu}_i - \bar{y}_i\left(t\right)\right) - \pi\left(y_i\right)\right]\mathrm{d}t + \sqrt{\alpha_i + \beta_i y_i\left(t\right)}\mathrm{d}\tilde{W}_i\left(t\right), \quad i = 1, 2, 3,$$

where $\pi\left(y_i\right)$ denotes the i-th element of the vector $\boldsymbol{\pi}\left(\boldsymbol{y}\right)$ in Eq. (7), and $\tilde{W}_i$ is a standard Brownian motion under the risk-neutral probability. We use the discretization step $\hat{\Delta} = \Delta/22$, where Δ is the discretization step used in the first and the second step of our estimation procedure Given Eqs. (8)-(11), the model-based volatility under the risk-neutral measure, at the j-th simulation, is:

$$\sigma^2_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}) = \hat{\sigma}^2_G + \frac{\sum_{i=1}^{3} \hat{s}^2_{i,T}\left(\hat{\alpha}_{i,T} + \hat{\beta}_{i,T}\hat{y}^{\lambda}_{i,t+k\hat{\Delta},h}\right)}{\tilde{s}^2_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda})}, \tag{54}$$

where

$$\tilde{s}_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}) = \hat{s}_{0,T} + \sum_{i=1}^{3} \hat{s}_{i,T}\hat{y}^{\lambda}_{i,t+k\hat{\Delta},h}, \tag{55}$$

and $\hat{\sigma}_{G,T}$ and $\hat{s}_{l,T}$ $l = 0, \cdots, 3$ are the standard deviation of stochastic secular growth and the reduced-form parameters obtained in the second step of the estimation procedure. Finally, we compute the simulated value of the model-based VIX, $\text{VIX}_{t,\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda})$, by integrating volatility over each month, as follows:

$$\text{VIX}_{t,\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}) = \sqrt{\frac{1}{\hat{\Delta}} \sum_{k=0}^{\hat{\Delta}^{-1}-1} \sigma^2_{t+(k+1)\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda})}. \tag{56}$$

By repeating the same procedure outlined above H times, we can then generate H paths of length $\mathcal{T}$. From now on, we simplify notation and index all parameter estimators and simulated factors by Δ, rather than $\hat{\Delta}$.

The estimator of $\hat{\boldsymbol{\lambda}}_{\mathcal{T}}$ in Eq. (26)

We have:

Proposition B3: *Under regularity conditions (Assumption B1 in Appendix B), if for some $\pi \in (0,1)$, $T, \mathcal{T}$, $\Delta\sqrt{T} \to 0$, $\Delta T \to \infty$,* and *$\mathcal{T}/T \to \pi$, then:*

$$\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\lambda}}_{\mathcal{T}} - \boldsymbol{\lambda}_0\right) \xrightarrow{d} \mathrm{N}\left(\mathbf{0}, \boldsymbol{V}_3\right), \quad \boldsymbol{V}_3 = \left(\boldsymbol{D}_3^\top \boldsymbol{D}_3\right)^{-1} \boldsymbol{D}_3^\top \left(\left(1 + \frac{1}{H}\right)(\boldsymbol{J}_3 - \boldsymbol{K}_3) + \boldsymbol{P}_3\right) \boldsymbol{D}_3 \left(\boldsymbol{D}_3^\top \boldsymbol{D}_3\right)^{-1},$$

where $\boldsymbol{\lambda}_0$ is as in Eq. (51), and the four matrices, $\boldsymbol{D}_3$, $\boldsymbol{J}_3$, $\boldsymbol{K}_3$ and $\boldsymbol{P}_3$, are defined in the proof below.

Proof: Given Assumptions B1(i) and B1(iii), for any $\boldsymbol{\lambda}$ in a compact set $\boldsymbol{\Lambda}_0$, $y^{\lambda}_{i,t+(k+1)\Delta,h}$, $i = 1, 2, 3$, $h = 1, \cdot\cdot\cdot, H$, is geometrically β-mixing, and has a stationary distribution with exponential tails. Thus, by Eqs. (54), (55) and (56), $\mathrm{VIX}_{t,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\phi}_0, \sigma_G, \boldsymbol{\lambda}_0\right)$ is also geometrically β-mixing with exponential tails. Therefore, $\mathrm{VIX}_{t,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\phi}_0, \sigma_G, \boldsymbol{\lambda}_0\right)$ has enough finite moments to satisfy sufficient conditions for the law of large numbers and the central limit theorem to apply. Next, note that $\mathrm{VIX}_{t,\Delta,h}\left(\boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_G, \boldsymbol{\lambda}\right)$ is continuously differentiable in the interior of $\boldsymbol{\Phi}_0 \times \boldsymbol{\Theta}_0 \times \boldsymbol{\Sigma}_G \times \boldsymbol{\Lambda}_0$ (for some compact set $\boldsymbol{\Sigma}_G$) and, hence, the uniform law of large numbers also applies. Similarly as in the proof of Propositions B2, we take into account the contribution of parameter estimation error, arising because the risk-neutral paths of the factors are generated using $\hat{\boldsymbol{\phi}}_T$,$\hat{\boldsymbol{\theta}}_T$ and $\hat{\sigma}_{G,T}$, instead of the unknown $\boldsymbol{\phi}_0$, $\boldsymbol{\theta}_0$ and σ_G.

By an argument similar to that in the proof of Proposition B1,

$$\begin{aligned}&\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\lambda}}_{\mathcal{T}} - \boldsymbol{\lambda}_0\right)\\ &= -\left(\boldsymbol{D}_3^\top \boldsymbol{D}_3\right)^{-1} \boldsymbol{D}_3^\top \left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}_0) - \boldsymbol{\psi}_0\right) - \sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}} - \boldsymbol{\psi}_0\right)\right) + o_p(1),\end{aligned}$$

where

$$\boldsymbol{D}_3 = \operatorname*{plim}_{\mathcal{T}\to\infty} \nabla_\lambda \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\boldsymbol{\phi}_0, \boldsymbol{\theta}_0, \sigma_G, \boldsymbol{\lambda}_0\right)\right),$$

and along the same lines as those in the proof of Proposition B2,

$$\begin{aligned}&\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}_0) - \boldsymbol{\psi}_0\right)\\ &= \sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\boldsymbol{\phi}_0, \boldsymbol{\theta}_0, \sigma_G, \boldsymbol{\lambda}_0\right) - \boldsymbol{\psi}_0\right) + \sqrt{\pi}\boldsymbol{F}_{\theta_0}^\top \sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right) + \sqrt{\pi}\boldsymbol{F}_{\phi_0}^\top \sqrt{T}\left(\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0\right)\\ &+ \sqrt{\pi}\boldsymbol{F}_{\sigma_G}^\top \sqrt{T}\left(\hat{\sigma}_{G,T} - \sigma_G\right) + o_p(1),\end{aligned}$$

where $\pi = \lim_{T,\mathcal{T}\to\infty} \mathcal{T}/T$, and, with $\bar{\boldsymbol{\theta}}_T$,$\bar{\boldsymbol{\phi}}_T$ and $\bar{\sigma}_{G,T}$ denoting convex combinations of $(\hat{\boldsymbol{\theta}}_T, \boldsymbol{\theta}_0)$, $(\hat{\boldsymbol{\phi}}_T, \boldsymbol{\phi}_0)$ and $(\hat{\boldsymbol{\sigma}}_{G,T}, \sigma_G)$, respectively,

$$\boldsymbol{F}_{\theta_0}^\top = \operatorname*{plim}_{T,\mathcal{T}\to\infty} \nabla_\theta \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\bar{\boldsymbol{\phi}}_T, \bar{\boldsymbol{\theta}}_T, \bar{\sigma}_{G,T}, \boldsymbol{\lambda}_0)\right),$$

with $\boldsymbol{F}_{\phi_0}^\top$ and $\boldsymbol{F}_{\sigma_G}^\top$ defined analogously. Therefore, by the same argument as those in the proofs of Propositions B1 and B2,

$$\mathrm{Avar}\left(\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\lambda}}_{\mathcal{T}} - \boldsymbol{\lambda}_0\right)\right) = \left(\boldsymbol{D}_3^\top \boldsymbol{D}_3\right)^{-1} \boldsymbol{D}_3^\top \left(\left(1 + \frac{1}{H}\right)(\boldsymbol{J}_3 - \boldsymbol{K}_3) + \boldsymbol{P}_3\right) \boldsymbol{D}_3 \left(\boldsymbol{D}_3^\top \boldsymbol{D}_3\right)^{-1},$$

where

$$\boldsymbol{J}_3 = \mathrm{Avar}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}} - \boldsymbol{\psi}_0\right)\right) = \mathrm{Avar}\left(\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\boldsymbol{\phi}_0, \boldsymbol{\theta}_0, \sigma_G, \boldsymbol{\lambda}_0\right) - \boldsymbol{\psi}_0\right)\right), \quad \text{for all } h,$$

$$\boldsymbol{K}_3 = \mathrm{Acov}\left(\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\psi}}_{\mathcal{T},1}\left(\boldsymbol{\phi}_0, \boldsymbol{\theta}_0, \sigma_G, \boldsymbol{\lambda}_0\right) - \boldsymbol{\psi}_0\right), \sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\psi}}_{\mathcal{T},2}\left(\boldsymbol{\phi}_0, \boldsymbol{\theta}_0, \sigma_G, \boldsymbol{\lambda}_0\right) - \boldsymbol{\psi}_0\right)^\top\right),$$

and

$$\begin{aligned}
\boldsymbol{P}_3 = & \ \pi \boldsymbol{F}_{\theta_0}^{\top} \mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right)\right) \boldsymbol{F}_{\theta_0} + \pi \boldsymbol{F}_{\phi_0}^{\top} \mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0\right)\right) \boldsymbol{F}_{\phi_0} \\
& + \pi \boldsymbol{F}_{\sigma_G}^{\top} \mathrm{Avar}\left(\sqrt{T}\left(\hat{\sigma}_{G,T} - \sigma_G\right)\right) \boldsymbol{F}_{\sigma_G} \\
& + 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{T,\Delta,h}\left(\boldsymbol{\phi}_0, \boldsymbol{\theta}_0, \sigma_G, \boldsymbol{\lambda}_0\right) - \boldsymbol{\psi}_0\right), \boldsymbol{F}_{\phi_0}^{\top} \sqrt{T}\left(\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0\right)\right) \\
& + 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{T,\Delta,h}\left(\boldsymbol{\phi}_0, \boldsymbol{\theta}_0, \boldsymbol{\lambda}_0\right) - \boldsymbol{\psi}_0\right), \boldsymbol{F}_{\theta_0}^{\top} \sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right)\right) \\
& + 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{T,\Delta,h}\left(\boldsymbol{\phi}_0, \boldsymbol{\theta}_0, \boldsymbol{\lambda}_0\right) - \boldsymbol{\psi}_0\right), \boldsymbol{F}_{\sigma_G}^{\top} \sqrt{T}\left(\hat{\sigma}_{G,T} - \sigma_G\right)\right) \\
& - 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\tilde{\boldsymbol{\psi}}_T - \boldsymbol{\psi}_0\right), \boldsymbol{F}_{\phi_0}^{\top} \sqrt{T}\left(\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0\right)\right) \\
& - 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\tilde{\boldsymbol{\psi}}_T - \boldsymbol{\psi}_0\right), \boldsymbol{F}_{\theta_0}^{\top} \sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right)\right) \\
& - 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\tilde{\boldsymbol{\psi}}_T - \boldsymbol{\psi}_0\right), \boldsymbol{F}_{\sigma_G}^{\top} \sqrt{T}\left(\hat{\sigma}_{G,T} - \sigma_G\right)\right).
\end{aligned}$$

B.2. Bootstrap estimates of the standard errors

We develop bootstrap standard errors consistent for $\boldsymbol{V}_1$, $\boldsymbol{V}_2$, and $\boldsymbol{V}_3$ of Propositions B1, B2, and B3. We draw B overlapping blocks of length l, with $T = Bl$, of

$$\boldsymbol{X}_t = (y_{1,t}, \cdots, y_{1,t-k_1}, y_{2,t}, \cdots, y_{2,t-k_2}, S_t, \cdots, S_{t-k_3}),$$

where k_1, k_2, k_3 depend on the lags we use in the auxiliary models. The resampled observations are:

$$\boldsymbol{X}_t^* = (y_{1,t}^*, \cdots, y_{1,t-k_1}^*, y_{2,t}^*, \cdots, y_{2,t-k_2}^*, S_t^*, \cdots, S_{t-k_3}^*).$$

Let P^* be the probability measure governing the resampled series, $\boldsymbol{X}_t^*$, and let $\mathrm{E}^*, \mathrm{var}^*$ denote the mean and the variance taken with respect to P^*, respectively. Further $O_p^*(1)$ and $o_p^*(1)$ denote, respectively, a term bounded in probability, and converging to zero in probability, under P^*, conditional on the sample and for all samples but a set of probability measure approaching zero.

For the implementation, we use block sizes of approximately $T^{1/4}$ and $T^{1/3}$ (see Lahiri (2003)), which give similar results. The standard errors reported in the main text are based on block sizes of approximately $T^{1/4}$. (Note: Whilst S_t^* does not not necessarily mimic the dependence of S_t, we just use S_t^* to compute R_t^* and Vol_t^*, which indeed mimic the dependence of R_t and Vol_t.)

Bootstrap Standard Errors for $\boldsymbol{\phi}$

The simulated samples for $y_{1,t}$ and $y_{2,t}$ are independent of the actual samples and are also independent across simulation replications. Also, as stated in Proposition B1, the estimators of the auxiliary model parameters, based on actual and simulated samples, have the same asymptotic variance. Hence, there is no need to resample the simulated series.

Given that the number of auxiliary model parameters and moment conditions is larger than the number of parameters to be estimated, we need to use an appropriate re-centering term. In the over-identified case, even if the population moment conditions have mean zero, the bootstrap moment conditions do not have mean zero, and a hence proper re-centering term is necessary (see, e.g., Hall and Horowitz (1996)).

Let $\tilde{\boldsymbol{\varphi}}_{T,i}^*$ be the bootstrap analog to $\tilde{\boldsymbol{\varphi}}_T$ at draw i, and define:

$$\hat{\boldsymbol{\phi}}_{T,i}^* = \arg\min_{\boldsymbol{\phi} \in \boldsymbol{\Phi}_0} \left\| \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi}) - \hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T) \right) - \left(\tilde{\boldsymbol{\varphi}}_{T,i}^* - \tilde{\boldsymbol{\varphi}}_T \right) \right) \right\|^2, \quad i = 1, \cdots, B.$$

We compute the bootstrap covariance matrix, as follows:

$$\hat{\boldsymbol{V}}_{1,T,B} = \frac{T}{B} \sum_{i=1}^{B} \left| \hat{\boldsymbol{\phi}}_{T,i}^* - \frac{1}{B} \sum_{i=1}^{B} \hat{\boldsymbol{\phi}}_{T,i}^* \right|_2 .$$

The next proposition shows that $\left(1 + \frac{1}{H}\right) \hat{\boldsymbol{V}}_{1,T,B}$, is a consistent estimator of $\boldsymbol{V}_1$, thereby allowing to compute asymptotically valid bootstrap standard errors.

Proposition B4: *Under the same assumptions of Proposition B1, if $l/T^{1/2} \to 0$ as $T, B, l \to \infty$, then for all $\varepsilon > 0$,*

$$\Pr\left(\omega : P^{*}\left(\left|\left(1+\frac{1}{H}\right)\hat{\boldsymbol{V}}_{1,T,B}-\boldsymbol{V}_{1}\right|>\varepsilon\right)\right)\to 0.$$

Proof: By the first order conditions and a mean value expansion around $\hat{\boldsymbol{\phi}}_T$,

$$\begin{aligned}
\mathbf{0} &= \nabla_{\phi}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T^{*})\right)^{\top}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\left(\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T^{*})-\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T)\right)-(\tilde{\boldsymbol{\varphi}}_T^{*}-\tilde{\boldsymbol{\varphi}}_T)\right)\\
&= \nabla_{\phi}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T^{*})\right)^{\top}(\tilde{\boldsymbol{\varphi}}_T-\tilde{\boldsymbol{\varphi}}_T^{*})\\
&\quad+\nabla_{\phi}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T^{*})\right)^{\top}\nabla_{\phi}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\bar{\boldsymbol{\phi}}_T^{*})\right)\left(\hat{\boldsymbol{\phi}}_T^{*}-\hat{\boldsymbol{\phi}}_T\right),
\end{aligned}$$

where $\bar{\boldsymbol{\phi}}_T^{*}$ is some convex combination of $\hat{\boldsymbol{\phi}}_T^{*}$ and $\hat{\boldsymbol{\phi}}_T$. Hence,

$$\begin{aligned}
&\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T^{*}-\hat{\boldsymbol{\phi}}_T\right)\\
&=\left(\nabla_{\phi}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,h}(\hat{\boldsymbol{\phi}}_T^{*})\right)^{\top}\nabla_{\phi}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,h}(\bar{\boldsymbol{\phi}}_T^{*})\right)\right)^{-1}\nabla_{\phi}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,h}(\hat{\boldsymbol{\phi}}_T^{*})\right)^{\top}\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T^{*}-\tilde{\boldsymbol{\varphi}}_T).
\end{aligned}$$

The Proposition follows, once we show that:

$$\mathrm{E}^{*}\left(\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T^{*}-\tilde{\boldsymbol{\varphi}}_T)\right)=o_p(1), \tag{57}$$

$$\mathrm{var}^{*}\left(\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T^{*}-\tilde{\boldsymbol{\varphi}}_T)\right)=\mathrm{var}\left(\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T-\boldsymbol{\varphi}_0)\right)+O_p(l/\sqrt{T}), \tag{58}$$

and for $\varepsilon > 0$,

$$\mathrm{E}^{*}\left(\left(\sqrt{T}\,\|\tilde{\boldsymbol{\varphi}}_T^{*}-\tilde{\boldsymbol{\varphi}}_T\|\right)^{2+\varepsilon}\right)=O_p(1). \tag{59}$$

Indeed, under conditions (57)-(58), we have that by the uniform law of large numbers, $\left|\nabla_{\phi}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,h}(\hat{\boldsymbol{\phi}}_T^{*})\right)-\boldsymbol{D}_1\right|=o_p^{*}(1)$. Hence,

$$\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T^{*}-\hat{\boldsymbol{\phi}}_T\right)=\left(\boldsymbol{D}_1^{\top}\boldsymbol{D}_1\right)^{-1}\boldsymbol{D}_1^{\top}\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T-\tilde{\boldsymbol{\varphi}}_T^{*})+o_p^{*}(1).$$

and, given (58), and recalling that $l/\sqrt{T}\to 0$,

$$\mathrm{var}^{*}\left(\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T^{*}-\tilde{\boldsymbol{\varphi}}_T)\right)=\mathrm{Avar}\left(\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T-\boldsymbol{\varphi}_0)\right)+o_p(1).$$

Given (59), the statement follows by Theorem 1 in Goncalves and White (2005).

Let us show (57), (58) and (59). We have,

$$\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T^{*}-\tilde{\boldsymbol{\varphi}}_T)=\sqrt{T}\left(\left(\tilde{\boldsymbol{\varphi}}_{1,T}^{*}-\tilde{\boldsymbol{\varphi}}_{1,T}\right),\left(\tilde{\boldsymbol{\varphi}}_{2,T}^{*}-\tilde{\boldsymbol{\varphi}}_{2,T}\right),\left(\bar{y}_1^{*}-\bar{y}_1\right),\left(\bar{y}_2^{*}-\bar{y}_2\right),\left(\hat{\sigma}_1^{*2}-\hat{\sigma}_1^{2}\right),\left(\hat{\sigma}_2^{*2}-\hat{\sigma}_2^{2}\right)\right)^{\top}.$$

Since each component of $\sqrt{T}\,(\tilde{\boldsymbol{\varphi}}_T^{*}-\tilde{\boldsymbol{\varphi}}_T)$ can be dealt with in the same way, we only consider $\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_{1,T}^{*}-\tilde{\boldsymbol{\varphi}}_{1,T}\right)$. Let $\boldsymbol{Y}_t$ be the vector containing all the regressors in Eq. (17), and $\boldsymbol{Y}_t^{*}$ be its bootstrap counterpart. By the first order conditions,

$$\begin{aligned}
\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_{1,T}^{*}-\tilde{\boldsymbol{\varphi}}_{1,T}\right)&=\left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t^{*}\boldsymbol{Y}_t^{*\top}\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=25}^{T}\boldsymbol{Y}_t^{*}\left(y_{1,t}^{*}-\boldsymbol{Y}_t^{*\top}\tilde{\boldsymbol{\varphi}}_{1,T}\right)\\
&=\left(\mathrm{E}(\boldsymbol{Y}_t\boldsymbol{Y}_t^{\top})\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=25}^{T}\boldsymbol{Y}_t^{*}\left(y_{1,t}^{*}-\boldsymbol{Y}_t^{*\top}\tilde{\boldsymbol{\varphi}}_{1,T}\right)+o_p^{*}(1),
\end{aligned}$$

as $\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t^*\boldsymbol{Y}_t^{*\top}-\mathrm{E}^*\left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t^*\boldsymbol{Y}_t^{*\top}\right)=o_p^*(1)$, and $\mathrm{E}^*\left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t^*\boldsymbol{Y}_t^{*\top}\right)=\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t\boldsymbol{Y}_t^\top+O_p(l/T)=\mathrm{E}\left(\boldsymbol{Y}_t\boldsymbol{Y}_t^\top\right)+o_p(1)$. We have,

$$\mathrm{E}^*\left(\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_{1,T}^*-\tilde{\boldsymbol{\varphi}}_{1,T}\right)\right)=\mathrm{E}(\boldsymbol{Y}_t\boldsymbol{Y}_t^\top)\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t\left(y_{1,t}-\boldsymbol{Y}_t^\top\tilde{\boldsymbol{\varphi}}_{1,T}\right)+O_p(l/\sqrt{T})=o_p(1).$$

This proves (57). Next,

$$\begin{aligned}
&\mathrm{var}^*\left(\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_{1,T}^*-\tilde{\boldsymbol{\varphi}}_{1,T}\right)\right)\\
=\;&\left(\mathrm{E}^*(\boldsymbol{Y}_t^*\boldsymbol{Y}_t^{*\top})\right)^{-1}\mathrm{var}^*\left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t^*\left(y_{1,t}^*-\boldsymbol{Y}_t^{*\top}\tilde{\boldsymbol{\varphi}}_{1,T}\right)\right)\left(\mathrm{E}^*(\boldsymbol{Y}_t^*\boldsymbol{Y}_t^{*\top})\right)^{-1}+o_p(1)\\
=\;&\left(\mathrm{E}(\boldsymbol{Y}_t\boldsymbol{Y}_t^\top)\right)^{-1}\left(\frac{1}{T}\sum_{j=-l}^{l}\sum_{t=25+l}^{T-l}\boldsymbol{Y}_t\boldsymbol{Y}_{t-j}^\top\tilde{\epsilon}_{1,t}\tilde{\epsilon}_{1,t-j}\right)\left(\mathrm{E}(\boldsymbol{Y}_t\boldsymbol{Y}_t^\top)\right)^{-1}+o_p(1)\\
=\;&\mathrm{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_{1,T}-\boldsymbol{\varphi}_{1,0}\right)\right)+o_p(1),
\end{aligned}$$

where $\tilde{\epsilon}_{1,t}=y_{1,t}-\boldsymbol{Y}_t^\top\tilde{\boldsymbol{\varphi}}_{1,T}$. This proves (58). Finally, as $\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t\boldsymbol{Y}_t^\top$ is full rank, for a generic constant C, and $\varepsilon>0$,

$$\mathrm{E}^*\left(\left(\sqrt{T}\left\|\tilde{\boldsymbol{\varphi}}_{1,T}^*-\tilde{\boldsymbol{\varphi}}_{1,T}\right\|\right)^{2+\varepsilon}\right)\leq C\mathrm{E}^*\left\|\frac{1}{\sqrt{T}}\sum_{t=25}^{T}\boldsymbol{Y}_t^{*\top}\left(y_{1,t}^*-\boldsymbol{Y}_t^{*\top}\tilde{\boldsymbol{\varphi}}_{1,T}\right)\right\|^{2+\varepsilon}.$$

By Lemma 2.1 in Goncalves and White (2005), $\mathrm{E}\left(\mathrm{E}^*\left\|\frac{1}{\sqrt{T}}\sum_{t=25}^{T}\boldsymbol{Y}_t^{*\top}\left(y_{1,t}^*-\boldsymbol{Y}_t^{*\top}\tilde{\boldsymbol{\varphi}}_{1,T}\right)\right\|^{2+\varepsilon}\right)=O(1)$. Hence, (59) follows by Markov inequality.

Bootstrap Standard Errors for $\boldsymbol{\theta}$

The model-based stock price series is simulated using the actual samples of the observable factors, and simulated samples for the unobservable factor and secular growth, $\ln G_{t,\Delta,h}^{\hat{\theta}_{G,T}}$. Thus, we need to take into account the contribution of $\boldsymbol{K}_2$, the covariance between simulated and sample paths, as well as the contribution of $\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T}-\boldsymbol{\theta}_G\right)$. To construct the resampled simulated stock prices through Eq. (20), we need to resample secular growth, $\ln G_{t,\Delta,h}^{\theta_G}$ through $\hat{\boldsymbol{\theta}}_{G,T}^*$, the boostrap analog to $\hat{\boldsymbol{\theta}}_{G,T}$. As secular growth is a geometric Brownian motion, we cannot use the block bootstrap to obtain $\hat{\boldsymbol{\theta}}_{G,T}^*$. Instead, we rely on the residual-based bootstrap of Paparoditis and Politis (2003). Let $\hat{\epsilon}_t=\left(\ln G_t-\ln G_{t-1}-\hat{g}_T+\frac{1}{2}\hat{\sigma}_{G,T}^2\right)/\hat{\sigma}_{G,T}$, where G_t is the secular growth, extracted through the Hodrick-Prescott filter, as discussed in the main text. Resample from $\hat{\epsilon}_t-\frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_t$, to obtain $\hat{\epsilon}_1^*,\cdots,\hat{\epsilon}_T^*$. Next, define

$$\ln G_t^*=\begin{cases}\ln G_1, & \text{for } t=1\\ \ln G_{t-1}^*+\hat{g}_T-\frac{1}{2}\hat{\sigma}_{G,T}^2+\hat{\sigma}_{G,T}\hat{\epsilon}_1^*, & \text{for } t=2,\cdots,T\end{cases}$$

Use $\ln G_t^*$ to get the bootstrap estimator, $\hat{\boldsymbol{\theta}}_{G,T}^*=\left(\hat{g}_T^*,\hat{\sigma}_{G,T}^*\right)$. Use Eq. (4), to generate $\ln G_{t,\Delta,h}^{\hat{\theta}_{G,T}^*}$, and resample blocks from it, to obtain $\ln G_{t,\Delta,h}^{*\hat{\theta}_{G,T}^*}$. Construct the resampled simulated stock price series as:

$$\ln S_{t,\Delta,h}^{*\theta}(\hat{\boldsymbol{\theta}}_{G,T}^*)=\ln G_{t,\Delta,h}^{*\hat{\theta}_{G,T}^*}+\ln\left(s_0+s_1y_{1,t}^*+s_2y_{2,t}^*+Z_{t,\Delta,h}^{\theta_u,*}\right), \tag{60}$$

where $Z_{t,\Delta,h}^{\theta_u,*}$ is resampled from the simulated unobservable process $Z_{t,\Delta,h}^{\theta_u}$, and use $S_{t,\Delta,h}^{*\theta}(\hat{\boldsymbol{\theta}}_{G,T}^*)$ to construct $R_{t,\Delta,h}^*(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{G,T}^*)$ and $\mathrm{Vol}_{t,\Delta,h}^*(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{G,T}^*)$. Define,

$$\tilde{\boldsymbol{\vartheta}}_T^*=\left(\tilde{\boldsymbol{\vartheta}}_{1,T}^*,\tilde{\boldsymbol{\vartheta}}_{2,T}^*,\bar{R}^*,\overline{\mathrm{Vol}}^*\right)^\top,$$

where $\tilde{\boldsymbol{\vartheta}}^*_{1,T}, \tilde{\boldsymbol{\vartheta}}^*_{2,T}$ are the estimators of the auxiliary models obtained using resampled observations, and $\overline{R}^*$, $\overline{\mathrm{Vol}}^*$ are the sample means of $R^*_t = \ln(S^*_t/S^*_{t-12})$ and $\mathrm{Vol}^*_t = \sqrt{6\pi} \cdot \frac{1}{12}\sum_{i=1}^{12} \left|\ln\left(S^*_{t+1-i}/S^*_{t-i}\right)\right|$, with S^*_t being the resampled series of the observable stock price process S_t, and

$$\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T}) = \left(\hat{\boldsymbol{\vartheta}}^*_{1,T,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T}), \hat{\boldsymbol{\vartheta}}^*_{2,T,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T}), \bar{R}^*_{\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T}), \overline{\mathrm{Vol}}^*_{\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T})\right)^\top,$$

where $\hat{\boldsymbol{\vartheta}}^*_{1,T,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T})$ and $\hat{\boldsymbol{\vartheta}}^*_{2,T,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T})$ are the parameters of the auxiliary models estimated using resampled simulated observations, and $\overline{R}^*_{\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T})$, $\overline{\mathrm{Vol}}^*_{\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T})$ are the sample means of $R^*_{t,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T})$ and $\mathrm{Vol}^*_{t,\Delta,h}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T})$. Define:

$$\hat{\boldsymbol{\theta}}^*_{T,i} = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \left\| \frac{1}{H}\sum_{h=1}^{H} \left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h,i}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T,i}) - \hat{\boldsymbol{\vartheta}}^{\Delta}_{T,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\theta}}_{G,T,i})\right) - \left(\tilde{\boldsymbol{\vartheta}}^*_{T,i} - \tilde{\boldsymbol{\vartheta}}_T\right) \right\|^2, \quad i = 1, \cdots, B,$$

where $\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h,i}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T,i})$ and $\tilde{\boldsymbol{\vartheta}}^*_{T,i}$ denote the values of $\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}\left(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^*_{G,T,i}\right)$ and $\tilde{\boldsymbol{\vartheta}}^*_T$ at the i-th bootstrap replication. The bootstrap covariance matrix is:

$$\hat{\boldsymbol{V}}_{2,T,B} = \frac{T}{B}\sum_{i=1}^{B} \left| \hat{\boldsymbol{\theta}}^*_{T,i} - \frac{1}{B}\sum_{i=1}^{B} \hat{\boldsymbol{\theta}}^*_{T,i} \right|_2 .$$

The next proposition shows that $\hat{\boldsymbol{V}}_{2,T,B}$ is a consistent estimator of $\boldsymbol{V}_2$, and can then be used to obtain asymptotically valid bootstrap standard errors.

Proposition B5: *Under the same assumptions of Proposition B2, if $l/T^{1/2} \to 0$ as $T, B, l \to \infty$, then, for all $\varepsilon > 0$,*

$$\Pr\left(\omega : P^*\left(\left|\hat{\boldsymbol{V}}_{2,T,B} - \boldsymbol{V}_2\right| > \varepsilon\right)\right) \to 0.$$

Proof: By a similar argument as that in the proof of Proposition B4,

$$\begin{aligned}
&\sqrt{T}\left(\hat{\boldsymbol{\theta}}^*_T - \hat{\boldsymbol{\theta}}_T\right)\\
&= -\left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} \boldsymbol{D}_2^\top \sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\theta}}^*_{G,T}) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\theta}}_{G,T})\right) - \left(\tilde{\boldsymbol{\vartheta}}^*_T - \tilde{\boldsymbol{\vartheta}}_T\right)\right) + o_{p^*}(1)\\
&= -\left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} \boldsymbol{D}_2^\top \left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\theta}}_{G,T}) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\theta}}_{G,T})\right) - \left(\tilde{\boldsymbol{\vartheta}}^*_T - \tilde{\boldsymbol{\vartheta}}_T\right)\right)\right.\\
&\left.+\boldsymbol{C}_2^\top \sqrt{T}\left(\hat{\boldsymbol{\theta}}^*_{G,T} - \hat{\boldsymbol{\theta}}_{G,T}\right)\right) + o_{p^*(1)}.
\end{aligned}$$

Moreover, along the lines of the proof of Proposition B4, we can show that

$$\mathrm{E}^*\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}^*_T - \hat{\boldsymbol{\theta}}_T\right)\right) = o_{p^*}(1),$$

and:

$$\begin{aligned}
&\mathrm{Var}^*\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}^*_T - \hat{\boldsymbol{\theta}}_T\right)\right)\\
=\;& \left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} \boldsymbol{D}_2^\top \mathrm{Var}^*\left(\left(\sqrt{T}\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\theta}}_{G,T}) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\theta}}_{G,T})\right)\right)\right.\\
&\left.-\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}^*_T - \tilde{\boldsymbol{\vartheta}}_T\right) + \boldsymbol{C}_2^\top \sqrt{T}\left(\hat{\boldsymbol{\theta}}^*_{G,T} - \hat{\boldsymbol{\theta}}_{G,T}\right)\right) \boldsymbol{D}_2\left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} + o_p(1)\\
=\;& \left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} \boldsymbol{D}_2^\top \mathrm{Avar}\left(\left(\sqrt{T}\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\left(\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G) - \boldsymbol{\vartheta}_0\right)\right)\right.\\
&\left.-\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right) + \boldsymbol{C}_2^\top \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T} - \boldsymbol{\theta}_G\right)\right) \boldsymbol{D}_2\left(\boldsymbol{D}_2^\top \boldsymbol{D}_2\right)^{-1} + o_p(1).
\end{aligned}$$

Hence, $\mathrm{Var}^*\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}^*_T - \hat{\boldsymbol{\theta}}_T\right)\right) = \mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\right)\right) + o_p(1)$.

Finally, under the parameter restrictions in Assumptions B1(i) and B1(iii), Minkowski's inequality ensures that

$$\mathrm{E}^*\left(\left(\sqrt{T}\left\|\hat{\boldsymbol{\theta}}_T^* - \hat{\boldsymbol{\theta}}_T\right\|\right)^{2+\varepsilon}\right) = O_p(1).$$

The statement then follows from Theorem 1 in Goncalves and White (2005).

Bootstrap Standard Errors for $\boldsymbol{\lambda}$

As mentioned in the main text, the model free VIX index series is available only from 1990 and so in the third step we have a sample of length $\mathcal{T}$, instead of length T. Thus, we need to resample $y_{1,t}, y_{2,t}, S_t$ and VIX_t from the shorter sample, using blocksize l and number of blocks B, so that $lB = \mathcal{T}$. Also, we need to resample the unobservable factor from a sample of length $\mathcal{T}$, at the parameter estimate of $\boldsymbol{\theta}_u$ obtained in the previous step, $\hat{y}_{3,t,\Delta,h}^{\theta_u}$ say. Let $\mathrm{VIX}_{t,\Delta,h}^*(\boldsymbol{y}_t^*; \hat{\boldsymbol{\phi}}_T^*, \hat{\boldsymbol{\theta}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda})$ be the resampled model-based VIX, according to Eq. (55). Finally, let

$$\tilde{\boldsymbol{\psi}}_{\mathcal{T}}^* = \left(\tilde{\boldsymbol{\psi}}_{1,\mathcal{T}}^*, \overline{\mathrm{VIX}}^*, \hat{\sigma}_{\mathrm{VIX}}^*\right)^\top,$$

where $\tilde{\boldsymbol{\psi}}_{1,\mathcal{T}}^*$ is the parameter vector for the auxiliary model, estimated using $y_{1,t}^*, y_{2,t}^*$, and VIX_t^*, with VIX_t^* being the resampled series of the model-free VIX, and $\overline{\mathrm{VIX}}^*, \hat{\sigma}_{\mathrm{VIX}}^*$ are the sample mean and standard deviation of VIX_t^*, and:

$$\begin{aligned}
&\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}^*(\hat{\boldsymbol{\theta}}_T^*, \hat{\boldsymbol{\phi}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda}) \\
= \;& \left(\hat{\boldsymbol{\psi}}_{1,\mathcal{T},\Delta,h}^*(\hat{\boldsymbol{\theta}}_T^*, \hat{\boldsymbol{\phi}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda}), \overline{\mathrm{VIX}}_{\Delta,h}^*(\hat{\boldsymbol{\theta}}_T^*, \hat{\boldsymbol{\phi}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda}), \tilde{\sigma}_{\Delta,h,\mathrm{VIX}}^*(\hat{\boldsymbol{\theta}}_T^*, \hat{\boldsymbol{\phi}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda})\right)^\top,
\end{aligned}$$

where $\hat{\boldsymbol{\psi}}_{1,\mathcal{T},\Delta,h}^*(\hat{\boldsymbol{\phi}}_T^*, \hat{\boldsymbol{\theta}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda})$ is the parameter vector for the auxiliary model, estimated using $y_{1,t}^*$, $y_{2,t}^*$, and $\overline{\mathrm{VIX}}_{\Delta,h}^*(\hat{\boldsymbol{\theta}}_T^*, \hat{\boldsymbol{\phi}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda})$ and $\tilde{\sigma}_{\Delta,h,\mathrm{VIX}}^*(\hat{\boldsymbol{\theta}}_T^*, \hat{\boldsymbol{\phi}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda})$ are the sample mean and standard deviation of $\mathrm{VIX}_{t,\Delta,h}^*(\hat{\boldsymbol{\phi}}_T^*, \hat{\boldsymbol{\theta}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda})$. Define,

$$\hat{\boldsymbol{\lambda}}_{\mathcal{T}}^* = \arg\min_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}_0}\left\|\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}^*(\hat{\boldsymbol{\phi}}_T^*, \hat{\boldsymbol{\theta}}_T^*, \hat{\sigma}_{G,T}^*, \boldsymbol{\lambda}) - \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \hat{\sigma}_{G,T}, \hat{\boldsymbol{\lambda}}_{\mathcal{T}})\right) - \left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}^* - \tilde{\boldsymbol{\psi}}_{\mathcal{T}}\right)\right)\right\|^2.$$

Construct the bootstrap covariance matrix, as

$$\hat{\boldsymbol{V}}_{3,\mathcal{T},B} = \frac{\mathcal{T}}{B}\sum_{i=1}^{B}\left|\hat{\boldsymbol{\lambda}}_{\mathcal{T},i}^* - \frac{1}{B}\sum_{i=1}^{B}\hat{\boldsymbol{\lambda}}_{\mathcal{T},i}^*\right|_2,$$

where $\hat{\boldsymbol{\lambda}}_{\mathcal{T},i}^*$ denotes the value of $\hat{\boldsymbol{\lambda}}_{\mathcal{T}}^*$ at the i-th bootstrap replication.

The next proposition is the counterpart to Propositions B4 and B5. It shows that $\hat{\boldsymbol{V}}_{3,\mathcal{T},B}$ is a consistent estimator of $\boldsymbol{V}_3$, and can then provide asymptotically valid bootstrap standard errors.

Proposition B6: *Under the same assumptions of Proposition B3, if* $l/\mathcal{T}^{1/2} \to 0$ *as* $T, \mathcal{T}, B, l \to \infty$, *then, for all* $\varepsilon > 0$,

$$\Pr\left(\omega : P^*\left(\left|\hat{\boldsymbol{V}}_{3,\mathcal{T},B} - \boldsymbol{V}_3\right| > \varepsilon\right)\right) \to 0.$$

Proof: Follows by arguments nearly identical to those in the proof of Proposition B5.

C. References for the Supplemental material

Ang, A. and J. Liu, 2004. "How to Discount Cashflows with Time-Varying Expected Returns," *Journal of Finance* 59, 2745-2783.

Angeletos, G.-M., G. Lorenzoni and A. Pavan, 2008. "Wall Street and Silicon Valley: A Delicate Interaction," Working paper, MIT.

Bekaert, G. and S. R. Grenadier, 2001. "Stock and Bond Pricing in an Affine Economy," Working paper, Columbia Business School.

Bernanke, B. S., M. Gertler and S. Gilchrist, 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework," in J.B. Taylor and M. Woodford (Eds.): *Handbook of Macroeconomics*, Vol. 1C, Chapter 21, p. 1341-1393.

Glasserman, P. and K.-K. Kim, 2010. "Moment Explosion and Stationary Distributions in Affine Diffusion Models," *Mathematical Finance* 20, 1-34.

Goncalves, S. and H. White, 2005. "Bootstrap Standard Error Estimates for Linear Regression," *Journal of the American Statistical Association* 100, 970-979.

Hall, P. and J. L. Horowitz, 1996. "Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators," *Econometrica* 64, 891-916.

Lahiri, S. N., 2003. *Resampling methods for dependent data*. Springer.

Mamaysky, H., 2002. "A Model for Pricing Stocks and Bonds," Working paper, Yale School of Management.

Paparoditis, E. and D.N. Politis, 2003. "Residual-Based Block Bootstrap for Unit Root Testing," *Econometrica* 71, 813-855.

Pardoux, E. and D. Talay, 1985. "Discretization and Simulation of Stochastic Differential Equations," *Acta Applicandae Mathematicae* 3, 23-47.