Macroeconomic Determinants of Stock Market Volatility and Volatility Risk-Premiums*

Valentina Corradi
University of Warwick

Walter Distaso
Imperial College Business School

Antonio Mele
London School of Economics

First draft: July 22, 2005. This version: October 6, 2010.

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 of stock market volatility cannot be merely explained by business cycle factors. Rather, it relates to the presence of some unobserved factor. At the same time, our model predicts that such an unobservable factor cannot explain the ups and downs stock volatility experiences over time—the "volatility of volatility." Instead, the volatility of stock 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.

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

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

*We thank Yacine Aït-Sahalia, Alessandro Beber, Bernard Dumas, Marcelo Fernandes, Christian Julliard, Mark Salmon, George Tauchen, Viktor Todorov, and seminar participants at EDHEC-Nice, European Central Bank, HEC Lausanne, IE Business School (Madrid), the Universities of Aarhus (CREATES), Goethe, Louvain and Warwick, the 2006 London-Oxford Financial Econometrics Study Group, the 2006 conference on Realized Volatility (Montréal), the 2008 CEMMAP conference at UCL, 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. Valentina Corradi and Walter Distaso gratefully acknowledge financial support from the ESRC under the grant RES-062-23-0311. Antonio Mele thanks the British EPSRC for financial support via grant EP/C522958/1. The usual disclaimer applies.

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 mainly interested in the direct effects time-varying 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 forecast 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 aggregate stock 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. We show a fundamental result: stock volatility could not be explained by macroeconomic factors only. 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. At the same time, our model reveals that the presence of macroeconomic factors is needed to explain the variability of stock volatility around its level—the volatility of aggregate stock 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 new and 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.

The second set of empirical results relates to the estimation of volatility-related risk-premiums. In broad terms, the volatility risk-premium is defined as the difference between the expectation of future stock market volatility under the risk-neutral and the true probability. It quantifies how much a representative agent is willing to pay to ensure that volatility will not 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 risk-adjusted expectation of future volatility, which is the new VIX index. The model successfully captures the dramatic movements in the VIX 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 future 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 markets 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, Gibson and Zhou (2004) and Bollerslev and Zhou (2006) 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 (2008), who develop a consumption-based rationale for the existence of the volatility risk-premium, 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 (2008) 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 these 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, then, 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 the very same factors, along the lines of Duffee (2002). We show that the resulting no-arbitrage stock price is affine in the factors.[1] Our model does not allow for jumps

[1] Our model differs from those in Bekaert and Grenadier (2001), Ang and Liu (2004) or Mamaysky (2002). For

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, Drechsler and Yaron (2008), Carr and Wu (2009), Todorov (2009), and Todorov and Tauchen (2009) 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, and can be understood as the outcome of two forces, which we need to tell apart from data: (i) the market participants' risk-aversion, and (ii) the dynamics of the fundamentals. We address this identification issue by exploiting derivatives data, related to variance swaps. The variance swap rate is, theoretically, the risk-adjusted expectation of the future integrated volatility within one month, and is calculated daily since 2003, and re-calculated back to 1990, by the CBOE, as 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. We implement this step by matching moments related to ex-post stock market returns, realized stock market volatility and the macroeconomic factors. In the third step, we use data on the new VIX index, and the macroeconomic factors, to estimate the risk-premiums parameters, by matching the impulse response function of the model-based VIX index to its empirical counterpart. The limiting distribution of our estimators is affected by parameter estimation error, arising because the estimators as of the last step depend on parameter estimates computed in previous steps. While we do characterize standard errors, theoretically, the actual computation of these errors is problematic, in practice. 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 illus-

example, 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.

trates the estimation strategy. Section 4 presents our empirical results. Section 5 concludes, and a technical appendix provides details omitted from the main text.

2 The model

2.1 The macroeconomic environment

We assume that a number of factors affect the development of aggregate macroeconomic variables. These factors form a vector-valued process $\boldsymbol{y}(t)$, solution to a n-dimensional affine diffusion,

$$\mathrm{d}\boldsymbol{y}(t) = \boldsymbol{\kappa}(\boldsymbol{\mu} - \boldsymbol{y}(t))\,\mathrm{d}t + \boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y}(t))\,\mathrm{d}\boldsymbol{W}(t), \tag{1}$$

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 $\boldsymbol{\beta}_i$. Appendix A reviews sufficient conditions that are known to ensure that Eq. (1) has a strong solution with $\boldsymbol{V}(\boldsymbol{y}(t))_{(ii)} > 0$ almost surely for all t.

While we do 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 at time 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. (1) by:

$$\ln(\mathrm{M}_j(t)/\mathrm{M}_j(t-12)) = f_j(\boldsymbol{y}(t)), \quad j = 1, \cdots, N_{\mathrm{M}}, \tag{2}$$

where the collection of functions $\{f_j\}$ determines how the factors dynamics impinge upon the observed macroeconomic variables. We now turn to model asset prices.

2.2 Risk-premiums and stock market volatility

We assume that asset prices are related to the vector of factors $\boldsymbol{y}(t)$ in Eq. (1), and that some of these factors affect developments in macroeconomic conditions, through Eq. (2). We assume that asset prices respond to movements in the factors affecting macroeconomic conditions.[2] Formally,

[2] 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 sphere of the economy.

we assume that there exists a rational pricing function $s(\boldsymbol{y}(t))$ such that the real stock price at time t, $s(t)$ say, is $s(t) \equiv s(\boldsymbol{y}(t))$. We let this price function be twice continuously differentiable in $\boldsymbol{y}$. By Itô's lemma, $s(t)$ satisfies,

$$\frac{\mathrm{d}s(t)}{s(t)} = m(\boldsymbol{y}(t), s(t))\,\mathrm{d}t + \frac{s_y(\boldsymbol{y}(t))^\top \boldsymbol{\Sigma V}(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))}\mathrm{d}\boldsymbol{W}(t), \tag{3}$$

where $s_y(\boldsymbol{y}) = [\frac{\partial}{\partial y_1}s(\boldsymbol{y}), \cdots, \frac{\partial}{\partial y_n}s(\boldsymbol{y})]^\top$ and m is a function we shall determine below by no-arbitrage conditions. By Eq. (3), the instantaneous variance of stock returns is

$$\sigma^2(t) \equiv \left\| \frac{s_y(\boldsymbol{y}(t))^\top \boldsymbol{\Sigma V}(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))} \right\|^2. \tag{4}$$

Next, 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 $\boldsymbol{W}(t)$, $t \leq T$, and P be the physical probability under which $\boldsymbol{W}(t)$ is defined. The Radon-Nikodym derivative of the risk-neutral probability Q with respect to P on $\mathbb{F}(T)$ is,

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

for some adapted risk-premium process $\boldsymbol{\Lambda}(t)$. We assume that each component of the risk-premium *process* $\Lambda^i(t)$ satisfies,

$$\Lambda^i(t) = \Lambda^i(\boldsymbol{y}(t)), \quad i = 1, \cdots, d,$$

for some *function* Λ^i. We also assume that the safe asset is elastically supplied such that the short-term rate r (say) is constant.[3]

Under the equivalent martingale measure, the stock price is solution to,

$$\frac{\mathrm{d}s(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))} = \left(r - \frac{\delta(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))}\right)\mathrm{d}t + \frac{s_y(\boldsymbol{y}(t))^\top \boldsymbol{\Sigma V}(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))}\mathrm{d}\hat{\boldsymbol{W}}(t), \tag{6}$$

where $\delta(\boldsymbol{y})$ is the instantaneous dividend rate, and $\hat{\boldsymbol{W}}$ is a Brownian motion defined under the risk-neutral probability Q.

2.3 No-arbitrage restrictions

There is obviously no freedom in modeling risk-premiums and stochastic volatility separately. Given a dividend process, volatility is uniquely determined, once we specify the risk-premiums.

[3] This assumption can be replaced with a weaker condition that the short-term rate is an affine function of the underlying state vector. In this case, Proposition 1 below would not hold, which might considerably hinder the actual estimation of the model.

Consider, then, the following "essentially affine" specification for the dynamics of the factors in Eq. (1), 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,

$$\boldsymbol{\Lambda}(\boldsymbol{y}) = \boldsymbol{V}(\boldsymbol{y})\boldsymbol{\lambda}_1 + \boldsymbol{V}^{-}(\boldsymbol{y})\boldsymbol{\lambda}_2\boldsymbol{y}, \tag{7}$$

for some d-dimensional vector $\boldsymbol{\lambda}_1$ and some $d \times n$ matrix $\boldsymbol{\lambda}_2$. The functional form for $\boldsymbol{\Lambda}$ is the same as that suggested by Duffee (2002) in the term-structure literature. If the matrix $\boldsymbol{\lambda}_2 = \boldsymbol{0}_{d\times n}$, then, $\boldsymbol{\Lambda}$ collapses to the standard "completely affine" specification introduced by Duffie and Kan (1996), in which the risk-premiums $\boldsymbol{\Lambda}$ are tied up to the volatility of the fundamentals, $\boldsymbol{V}(\boldsymbol{y})$. While it is reasonable to assume that risk-premiums are related to the *volatility* of fundamentals, the specification in Eq. (7) is more general, as it allows risk-premiums to be related to the *level* of the fundamentals, through the additional term $\boldsymbol{\lambda}_2\boldsymbol{y}$.

Finally, we determine the no-arbitrage stock price. Under regularity conditions (see Appendix A), and in the absence of bubbles, Eq. (6) implies that the stock price is,

$$s(\boldsymbol{y}) = \mathbb{E}\left[\int_0^\infty e^{-rt}\delta(\boldsymbol{y}(t))\,\mathrm{d}t \,\middle|\, \boldsymbol{y}(0) = \boldsymbol{y}\right], \tag{8}$$

where $\mathbb{E}$ is the expectation taken under the risk-neutral probability Q. We are only left with specifying how the instantaneous dividend process relates to the state vector $\boldsymbol{y}$. As it turns out, the previous assumption on the pricing kernel and the assumption that $\delta(\cdot)$ is affine in $\boldsymbol{y}$ implies that the stock price is also affine in $\boldsymbol{y}$. Precisely, let

$$\delta(\boldsymbol{y}) = \delta_0 + \boldsymbol{\delta}^\top \boldsymbol{y}, \tag{9}$$

for some scalar δ_0 and some vector $\boldsymbol{\delta}$.[4] We have:

Proposition 1: *Let the risk-premiums and the instantaneous dividend rate be as in Eqs. (7) and (9). Then, under a technical regularity condition in Appendix A (condition (A2)), we have that: (i) Eq. (8) holds; and (ii) the rational stock price function $s(\boldsymbol{y})$ is linear in the state vector $\boldsymbol{y}$, viz*

$$s(\boldsymbol{y}) = \frac{\delta_0 + \boldsymbol{\delta}^\top (\boldsymbol{D} + r\boldsymbol{I}_{n\times n})^{-1}\boldsymbol{c}}{r} + \boldsymbol{\delta}^\top (\boldsymbol{D} + r\boldsymbol{I}_{n\times n})^{-1}\boldsymbol{y}, \tag{10}$$

[4] Eq. (9) makes the dividend stationary as soon as $\boldsymbol{y}(t)$ is stationary. Alternatively, we might assume that the dividend as of time t is $e^{gt}\delta(\boldsymbol{y}(t))$, for some constant g, where $\delta(\cdot)$ is as in Eq. (9). In this case, the price function is given by $e^{gt}s(\boldsymbol{y})$, where $s(\cdot)$ is the price function in Proposition 1, with r replaced by $r-g$. Such a more general formulation with a deterministic *trend* for the real stock price, would not alter our results in the empirical section, as our estimators do not rely on the assumption of absence of such a trend.

where

$$\boldsymbol{c} = \boldsymbol{\kappa}\boldsymbol{\mu} - \boldsymbol{\Sigma} \begin{pmatrix} \alpha_1 \lambda_{1(1)} & \cdots & \alpha_d \lambda_{1(d)} \end{pmatrix}^\top \tag{11}$$

$$\boldsymbol{D} = \boldsymbol{\kappa} + \boldsymbol{\Sigma} \left[\begin{pmatrix} \lambda_{1(1)} \boldsymbol{\beta}_1^\top & \cdots & \lambda_{1(d)} \boldsymbol{\beta}_d^\top \end{pmatrix}^\top + \boldsymbol{I}^- \boldsymbol{\lambda}_2 \right], \tag{12}$$

$\boldsymbol{I}^-$ *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* $\boldsymbol{\lambda}_1$.

Proposition 1 allows us to single out the no-arbitrage restrictions between stochastic volatility and risk-premiums. By Eq. (4), and the expression for the stock price in Eq. (10), we have:

$$\sigma(\boldsymbol{y}(t)) \equiv \sigma(t) = \frac{\sqrt{\left\| \boldsymbol{\delta}^\top (\boldsymbol{D} + r\boldsymbol{I}_{n\times n})^{-1} \boldsymbol{\Sigma} \boldsymbol{V}(\boldsymbol{y}(t)) \right\|^2}}{\frac{\delta_0 + \boldsymbol{\delta}^\top (\boldsymbol{D} + r\boldsymbol{I}_{n\times n})^{-1} \boldsymbol{c}}{r} + \boldsymbol{\delta}^\top (\boldsymbol{D} + r\boldsymbol{I}_{n\times n})^{-1} \boldsymbol{y}(t)}. \tag{13}$$

This expression for the stock volatility clarifies why our approach is distinct from that in the standard stochastic volatility literature. In this literature, the asset price and, hence, its volatility, is taken as given, and volatility and volatility risk-premiums are modeled independently of each other. For example, in the celebrated Heston's (1993) model, the stock price is solution to,

$$\begin{cases} \dfrac{\mathrm{d}s(t)}{s(t)} = m_H(t)\,\mathrm{d}t + v(t)\,\mathrm{d}W_1(t) \\ \mathrm{d}v^2(t) = \kappa_v \left(\mu_v - v^2(t)\right)\mathrm{d}t + \sigma_v v(t) \left(\rho \mathrm{d}W_1(t) + \sqrt{1-\rho^2}\mathrm{d}W_2(t)\right) \end{cases} \tag{14}$$

for some adapted process $m_H(t)$ and some constants $\kappa_v, \mu_v, \sigma_v, \rho$. In this model, the volatility risk-premium is specified separately from the volatility process. Many empirical studies have followed the lead of this model (e.g., Chernov and Ghysels (2000), Corradi and Distaso (2006), Garcia, Lewis, Pastorello and Renault (2007)). Moreover, a recent focus in this empirical literature is to examine how the risk-compensation for stochastic volatility is related to the business cycle (e.g., Bollerslev, Gibson and Zhou (2004)). While the empirical results in these papers are ground breaking, the Heston's model does not predict that there is any relation between stochastic volatility, volatility risk-premiums and the business cycle.

Our model works differently, as it places restrictions on the asset price process directly, through our assumptions about the fundamentals of the economy, i.e. the dividend process in Eq. (9) and the risk-premiums in Eq. (7). In our model, it is the asset price process that determines, endogenously, the volatility dynamics. For this reason, the model predicts that stock volatility embeds information about risk-corrections that agents require to invest in the stock market, as Eq. (13) makes clear. We shall make use of this observation in the empirical part of the paper. We now turn to describe which measure of stock volatility we use to proceed with such a critical step of our analysis.

2.4 Arrow-Debreu adjusted volatility

In September 2003, the Chicago Board Option Exchange (CBOE) changed its volatility index VIX to approximate the variance swap rate of the S&P 500 Compounded index. The new index reflects recent advances in the option pricing literature. Given an asset price process $s(t)$ that is continuous in time (as for the asset price of our model in Eq. (10)), and all available information $\mathbb{F}(t)$ at time t, consider the economic value of the future integrated variance on a given interval $[t, T]$, which is, approximately, the sum of the future variance weighted with the Arrow-Debreu state prices:

$$IV_{t,T} = \int_t^T \mathbb{E}\left[\left(\frac{\mathrm{d}}{\mathrm{d}\tau} \mathrm{var}\left[\ln s(\tau) \middle| \mathbb{F}(u) \right] \Big|_{\tau=u} \right) \middle| \mathbb{F}(t) \right] \mathrm{d}u. \tag{15}$$

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:

$$\mathbb{E}\left[IV_{t,T} \middle| \mathbb{F}(t) \right] = 2e^{-r(T-t)} \left[\int_0^{F(t)} \frac{P(t,T,K)}{K^2} \mathrm{d}K + \int_{F(t)}^{\infty} \frac{C(t,T,K)}{K^2} \mathrm{d}K \right], \tag{16}$$

where $F(t) = e^{r(T-t)} s(t)$ is the forward price, and $C(t,T,K)$ and $P(t,T,K)$ are the prices as of time t of a call and a put option expiring at T and struck at K. A variance swap is a contract with payoff proportional to the difference between the realized integrated variance, (15), and some strike price, the variance swap rate. In the absence of arbitrage opportunities, the variance swap rate is given by Eq. (16).

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:

$$\mathbb{E}\left[IV_{t,T} \middle| \boldsymbol{y}(t) = \boldsymbol{y} \right] = \int_t^T \mathbb{E}\left[\sigma^2(u) \middle| \boldsymbol{y}(t) = \boldsymbol{y} \right] \mathrm{d}u, \tag{17}$$

where $\sigma^2(t)$ is given in Eq. (13). An important task of this paper is to estimate the model so that it predicts a theoretical pattern of the VIX index that matches its empirical counterpart, computed by the CBOE through Eq. (16). Finally, note that our model makes predictions about future expected volatility under both the risk-neutral and the physical probability, P. Let, then, E denote the expectation taken under P. Our model allows to trace how the volatility risk-premium, defined as,

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

changes with changes in the factors $\boldsymbol{y}(t)$ in Eq. (1).

2.5 The leading model

We formulate a few specific assumptions to make the model amenable to empirical work. First, we assume that two macroeconomic aggregates, inflation and industrial production growth, are the only observable factors (say y_1 and y_2) affecting stock market developments. We define these factors as follows, $\ln(\mathrm{M}_j(t)/\mathrm{M}_j(t-12)) = \ln y_j(t)$, $j = 1, 2$, where $\mathrm{M}_1(t)$ is the consumer price index as of month t and $\mathrm{M}_2(t)$ is the industrial production as of month t. Hence, in terms of Eq. (2), the functions $f_j(\boldsymbol{y}) \equiv \ln y_j$. In Section 4.1, we discuss further the role these two macroeconomic factors have played in asset pricing. Second, we assume that a third unobservable factor y_3 affects the stock price, but not the two macroeconomic aggregates M_1 and M_2. Third, we consider a model in which the two macroeconomic factors y_1 and y_2 do not affect the unobservable factor y_3, although we allow for simultaneous feedback effects between inflation and industrial production growth. Therefore, we set, in Eq. (1),

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

where 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. Moreover, we take $\boldsymbol{\Sigma} = \boldsymbol{I}_{3\times3}$ and the vectors $\boldsymbol{\beta}_i$ so as to make y_j solution to,

$$\mathrm{d}y_j(t) = \left[\kappa_j\left(\mu_j - y_j(t)\right) + \bar{\kappa}_j\left(\bar{\mu}_j - \bar{y}_j(t)\right)\right]\mathrm{d}t + \sqrt{\alpha_j + \beta_j y_j(t)}\,\mathrm{d}W_j(t), \quad j = 1, 2, 3, \tag{18}$$

where, for brevity, we have set $\bar{\mu}_1 \equiv \mu_2$, $\bar{y}_1(t) \equiv y_2(t)$, $\bar{\mu}_2 \equiv \mu_1$, $\bar{y}_2(t) = y_1(t)$, $\bar{\kappa}_3 \equiv \bar{\mu}_3 \equiv \bar{y}_3(t) \equiv 0$ and, finally, $\beta_j \equiv \beta_{jj}$. We assume that, for each i, $\Pr\{V(\boldsymbol{y}(t))_{(ii)} > 0 \text{ all } t\} = 1$, which it does under the conditions reviewed in Appendix A.

We assume that the risk-premium process $\boldsymbol{\Lambda}$ satisfies the "essentially affine" specification in Eq. (7), where we take the matrix $\boldsymbol{\lambda}_2$ to be diagonal with diagonal elements equal to $\lambda_{2(j)} \equiv \lambda_{2(jj)}$, $j = 1, 2, 3$. The implication is that the *total* risk-premiums process defined as,

$$\boldsymbol{\pi}(\boldsymbol{y}) \equiv \boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})\boldsymbol{\Lambda}(\boldsymbol{y}) = \begin{pmatrix} \alpha_1\lambda_{1(1)} + \left(\beta_1\lambda_{1(1)} + \lambda_{2(1)}\right)y_1 \\ \alpha_2\lambda_{1(2)} + \left(\beta_2\lambda_{1(2)} + \lambda_{2(2)}\right)y_2 \\ \alpha_3\lambda_{1(3)} + \left(\beta_3\lambda_{1(3)} + \lambda_{2(3)}\right)y_3 \end{pmatrix} \tag{19}$$

depends on the factor y_j not only through the channel of the volatility of these factors (i.e. through the parameters β_{jj}), but also through the additional risk-premiums parameters $\lambda_{2(j)}$.

Finally, the instantaneous dividend process $\delta(t)$ in Eq. (9) satisfies,

$$\delta(\boldsymbol{y}) = \delta_0 + \delta_1 y_1 + \delta_2 y_2 + \delta_3 y_3. \tag{20}$$

Under these conditions, the asset price in Proposition 1 is given by,

$$s(\boldsymbol{y}) = s_0 + \sum_{j=1}^{3} s_j y_j, \tag{21}$$

where

$$\begin{aligned} s_0 &= \frac{1}{r}\left[\delta_0 + \sum_{j=1}^{3} s_j\left(\kappa_j\mu_j + \bar{\kappa}_j\bar{\mu}_j - \alpha_j\lambda_{1(j)}\right)\right], && (22)\\ s_j &= \frac{\delta_j\left(r+\kappa_i+\lambda_{1(i)}\beta_i+\lambda_{2(i)}\right)-\delta_i\bar{\kappa}_i}{\prod_{h=1}^{2}\left(r+\kappa_h+\lambda_{1(h)}\beta_h+\lambda_{2(h)}\right)-\bar{\kappa}_1\bar{\kappa}_2}, \quad \text{for } j,i\in\{1,2\} \text{ and } i\neq j, && (23)\\ s_3 &= \frac{\delta_3}{r+\kappa_3+\lambda_{1(3)}\beta_3+\lambda_{2(3)}}, && (24) \end{aligned}$$

and where $\bar{\kappa}_j$ and $\bar{\mu}_j$ are as in Eq. (18).

Note, then, an important feature of the model. The parameters $\lambda_{(1)i}$ and $\lambda_{(2)i}$ and δ_i cannot be identified from data on the asset price and the macroeconomic factors. Intuitively, the parameters $\lambda_{(1)i}$ and $\lambda_{(2)i}$ determine how sensitive the total risk-premium in Eq. (19) is to changes in the state process $\boldsymbol{y}$. Instead, the parameters δ_i determine how sensitive the dividend process in Eq. (20) is to changes in $\boldsymbol{y}$. Two price processes might be made observationally equivalent through judicious choices of the risk-compensation required to bear the asset or the payoff process promised by this asset (the dividend). The next section explains how to exploit the Arrow-Debreu adjusted volatility introduced in Section 2.4 to identify these parameters.

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 reduced-form parameters that link the equilibrium stock price to the three factors in Eq. (21), and the parameters of the process for the unobserved factor, $\boldsymbol{\theta}^\top = (\kappa_3, \mu_3, \alpha_3, \beta_3, s_0, s_j, j = 1, 2, 3)$, 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 time series behavior of 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, then, relies on an hybrid of Indirect Inference (Gouriéroux, Monfort and Renault, 1993) and the Simulated Generalized Method of Moments (Duffie and Singleton, 1993).[5]

[5] 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

3.1 Moment conditions for the macroeconomic factors

To simulate the factor dynamics in Eq. (18), 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 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 the parameter vector $\boldsymbol{\phi}$. Then, we estimate the following autoregressive models on both historical and simulated data, for $i = 1, 2$,

$$y_{i,t} = a^y_i + \sum_{j\in\{12,24\}} b^y_{i,1,j} y_{1,t-j} + \sum_{j\in\{12,24\}} b^y_{i,2,j} y_{2,t-j} + \epsilon^y_{i,t}, \tag{25}$$

and

$$y^{\phi}_{i,t,\Delta,h} = a^y_{i,h} + \sum_{j\in\{12,24\}} b^y_{i,1,j,h} y^{\phi}_{1,t-j,\Delta,h} + \sum_{j\in\{12,24\}} b^y_{i,2,j,h} y^{\phi}_{2,t-j,\Delta,h} + \epsilon^y_{i,t,h}. \tag{26}$$

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 estimators of the parameters in Eq. (25) for $i = 1, 2$, and $\bar{y}_i$ and $\hat{\sigma}_i$ are the sample mean and standard deviation of the macroeconomic factors. Likewise, define $\hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi})$ to be the simulated counterpart to $\tilde{\boldsymbol{\varphi}}_T$ at simulation h, including the ordinary least squares estimator of the parameters in Eq. (26), and the sample means and standard deviations of the macroeconomic factors.

The estimator of $\boldsymbol{\phi}$, the parameters of the process underlying the macroeconomic factors, is:

$$\hat{\boldsymbol{\phi}}_T \equiv \arg\min_{\boldsymbol{\phi}\in\boldsymbol{\Phi}_0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h}(\boldsymbol{\phi}) - \tilde{\boldsymbol{\varphi}}_T \right\|^2, \tag{27}$$

where $\boldsymbol{\Phi}_0$ is a compact set of $\boldsymbol{\Phi}$, a parameter set defined in Appendix B.1. Naturally, this estimator of $\boldsymbol{\phi}$, analyzed in Proposition 2 below (as well as those of $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$ in Propositions 3 and 4), depends on the discretization interval, Δ, although we do not make this dependence explicit, to alleviate notation.

We have:

Proposition 2: *Under regularity conditions (Assumption B1(i)-(iii) in Appendix B), as $T \to \infty$ and $\Delta\sqrt{T} \to 0$,*

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

(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.

where the two matrices $\boldsymbol{D}_1$ and $\boldsymbol{J}_1$ are defined in Appendix B.1, and $\boldsymbol{\phi}_0$ is the minimizer of the moment conditions in Eq. (27) for $T \to \infty$ and $\Delta \to 0$.

3.2 Moment conditions for realized returns and volatility

Data on macroeconomic factors and stock returns do not allow us to identify all the structural parameters of the model: the parameters s_j in Eq. (21) depend on the structural parameters, as Eqs. (22)-(24) show. In particular, we cannot identify the parameters related to the dividend process and the risk-premiums parameters: there are many combinations of $\boldsymbol{\delta}$ and $\boldsymbol{\lambda}$ giving rise to the same stock price. In this second step, we estimate the reduced-form parameters, s_j, and the parameters of the process for the unobservable factor y_3, $(\kappa_3, \mu_3, \alpha_3, \beta_3)$. The parameters $\boldsymbol{\lambda}$ shall be identified, and estimated, in a third and final step, described in the next section.

Even proceeding in this way, we are not able to tell apart the loading on the unobservable factor, s_3, from the parameters underlying the dynamics of the very same unobservable process, $(\kappa_3, \mu_3, \alpha_3, \beta_3)$, as this factor is independent of the observable ones. To address this issue, we impose the normalization $\mu_3 \equiv 1$, and define a new factor $Z(t) = s_3 y_3(t)$, which has dynamics:

$$\mathrm{d}Z(t) = \kappa_3 \left(s_3 - Z(t)\right) \mathrm{d}t + \sqrt{B + CZ(t)}\mathrm{d}W_3\left(t\right),$$

where $B = \alpha_3 s_3^2$ and $C = \beta_3 s_3$. We simulate H paths of length T of the unobservable factor $Z\left(t\right)$, using a Milstein approximation with discrete interval Δ, and sample it at the same frequency as the data, obtaining for $\boldsymbol{\theta}_u = (\kappa_3, \alpha_3, \beta_3, s_3)$ and simulation h, the series $Z_{t,\Delta,h}^{\theta_u}$. Likewise, let $s_{t,\Delta,h}^{\theta}$ be the simulated series of the stock price, when the parameters are fixed at $\boldsymbol{\theta}$:

$$s_{t,\Delta,h}^{\theta} = s_0 + s_1 y_{1,t} + s_2 y_{2,t} + Z_{t,\Delta,h}^{\theta_u}, \tag{28}$$

where we fix the intercept at $s_0 = \bar{s} - s_1\bar{y}_1 - s_2\bar{y}_2 - s_3$, and where $\bar{s}$, $\bar{y}_1$, and $\bar{y}_2$ are the sample means of the observed stock price index, S_t say, and 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:

$$\mathrm{Vol}_t = \sqrt{6\pi} \cdot \frac{1}{12} \sum_{i=1}^{12} \left| \ln\left(\frac{S_{t+1-i}}{S_{t-i}}\right) \right|. \tag{29}$$

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

Our estimator relies on the following two auxiliary models:

$$R_t = a^{\mathrm{R}} + b_1^{\mathrm{R}} y_{1,t-12} + b_2^{\mathrm{R}} y_{2,t-12} + \epsilon_t^{\mathrm{R}}, \tag{30}$$

and

$$\mathrm{Vol}_t = a^{\mathrm{V}} + \sum_{i\in\{6,12,18,24,36,48\}} b_i^{\mathrm{V}} \mathrm{Vol}_{t-i} + \sum_{i\in\{12,24,36,48\}} b_{1,i}^{\mathrm{V}} y_{1,t-i} + \sum_{i\in\{12,24,36,48\}} b_{2,i}^{\mathrm{V}} y_{2,t-i} + \epsilon_t^{\mathrm{V}}. \quad (31)$$

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 returns and volatility, $\tilde{\boldsymbol{\vartheta}}_{1,T}$ is the ordinary least squares estimate of the parameters in Eq. (30) and $\tilde{\boldsymbol{\vartheta}}_{2,T}$ is the ordinary least squares estimate of the parameters in Eq. (31). Let $\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\boldsymbol{\theta})$ be the simulated counterpart to $\tilde{\boldsymbol{\vartheta}}_T$ at simulation h.

The estimator of $\boldsymbol{\theta}$, the vector including the reduced-form parameters s_j and the parameters related to process of the unobservable factor, is:

$$\hat{\boldsymbol{\theta}}_T = \arg\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}_0} \left\| \frac{1}{H} \sum_{h=1}^{H} \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\boldsymbol{\theta}) - \tilde{\boldsymbol{\vartheta}}_T \right\|^2, \tag{32}$$

where $\boldsymbol{\Theta}_0$ is a compact set of $\boldsymbol{\Theta}$, a parameter set defined in Appendix B.1.

We have:

Proposition 3: *Under regularity conditions (Assumption B1(i)-(iv) in Appendix B), as $T \to \infty$ and $\Delta\sqrt{T} \to 0$,*

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

where the three matrices $\boldsymbol{D}_2$, $\boldsymbol{J}_2$ and $\boldsymbol{K}_2$ are defined in Appendix B.1, and $\boldsymbol{\theta}_0$ is the minimizer of the moment conditions in Eq. (32) for $T \to \infty$ and $\Delta \to 0$.

Note that the structure of the asymptotic covariance matrix is different from that in Proposition 2. The difference is the presence of the matrix $\boldsymbol{K}_2$, which captures the covariance across paths at different simulation replications, as well as the covariance between actual and simulated paths. Indeed, we are simulating the stock price process, conditionally upon the sample realizations for the observable factors, thus performing conditional simulated inference. This feature of the method results in a correlation between the auxiliary parameter estimates obtained over all the simulations. It is immediate to see that the use of observed values of $y_{1,t}$ and $y_{2,t}$ in (28), provides an efficiency improvement over unconditional (simulated) inference.

3.3 Estimation of the risk-premium parameters

Sample data on the macroeconomic factors and stock prices do not suffice to identify the risk-premium parameters, $\boldsymbol{\lambda}$. We identify, and estimate, $\boldsymbol{\lambda}$ by matching moments and impulse response functions of the model-based VIX to those of the model free VIX index.

The VIX, as defined in Eq. (16), is available only from 1990. Hence, in this stage, we use a sample of $\mathcal{T}$ observations, with $\mathcal{T} < T$. As for the theoretical counterpart to the VIX, consider the instantaneous stock volatility predicted by the model, as defined in Eq. (13), $\sigma(\boldsymbol{y}(t))$. The VIX index predicted by the model is,

$$\mathrm{VIX}(\boldsymbol{y}(t)) \equiv \sqrt{\frac{1}{T-t}\int_t^T \mathbb{E}[\sigma^2(\boldsymbol{y}(u)) \mid \boldsymbol{y}(t)=\boldsymbol{y}]\mathrm{d}u}, \tag{33}$$

where $T-t$ equals one month, and $\mathbb{E}$ is the expectation under the risk-neutral probability. Although $\mathrm{VIX}(\boldsymbol{y})$ is not known in closed-form, it can be accurately approximated through simulations. Appendix B.1 provides details on these simulations, upon which we base the proofs of Proposition 4 below. Note, also, that in the actual computation of Eq. (33), we replace the unknown parameters $s_0, s_j, \kappa_j, \alpha_j, \beta_j$ $j=1,2,3$ and $\bar{\kappa}_i, \mu_i$, $i=1,2$ with their estimated counterparts computed in the previous two steps: $\hat{\boldsymbol{\theta}}_T$ and $\hat{\boldsymbol{\phi}}_T$. This leads to parameter estimation error, which we shall have to take into account. As in the previous stage of the estimation, we make use of the observed samples for the macroeconomic factors $y_{1,t}, y_{2,t}$, and simulate samples for the latent factor only. Note, finally, that given $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, we can now identify $\boldsymbol{\lambda}$ from $\boldsymbol{c}$ and $\boldsymbol{D}$.

In the sequel, we rely on the following auxiliary model:

$$\mathrm{VIX}_t = a^{\mathrm{VIX}} + b^{\mathrm{VIX}}\mathrm{VIX}_{t-1} + \sum_{i\in\{36,48\}} b_{1,i}^{\mathrm{VIX}} y_{1,t-i} + \sum_{i\in\{36,48\}} b_{2,i}^{\mathrm{VIX}} y_{2,t-i} + \epsilon_t^{\mathrm{VIX}}. \tag{34}$$

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 ordinary least squares estimator of the parameters in Eq. (34), and $\overline{\mathrm{VIX}}$ and $\hat{\sigma}_{\mathrm{VIX}}$ are the sample mean and standard deviation of the VIX index. Likewise, define $\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda})$, the simulated counterpart to $\tilde{\boldsymbol{\psi}}_{\mathcal{T}}$ at simulation h, obtained through simulations of the model-implied index $\mathrm{VIX}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda})$, where the two macroeconomic factors, $y_{1,t}$ and $y_{2,t}$, are fixed at their sample values.

The estimator of $\boldsymbol{\lambda}$, the parameters underlying the risk-premium process, is:

$$\hat{\boldsymbol{\lambda}}_{\mathcal{T}} = \arg\min_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}_0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \boldsymbol{\lambda}) - \tilde{\boldsymbol{\psi}}_{\mathcal{T}} \right\|^2, \tag{35}$$

for some compact set $\boldsymbol{\Lambda}_0$.

We have:

Proposition 4: *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:*

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

where the four matrices $\boldsymbol{D}_3$, $\boldsymbol{J}_3$, $\boldsymbol{K}_3$ and $\boldsymbol{P}_3$ are defined in Appendix B.1, and $\boldsymbol{\lambda}_0$ is the minimizer of the moment conditions in Eq. (35) for $\Delta\sqrt{T} \to 0$, $\Delta T \to \infty$.

Note that the matrix $\boldsymbol{P}_3$ captures the contribution of parameter estimation error. The estimation error arises because the model-implied VIX index, $\mathrm{VIX}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda})$, is simulated using parameters estimated in the previous two stages, $\hat{\boldsymbol{\phi}}_T$ and $\hat{\boldsymbol{\theta}}_T$.

3.4 Bootstrap Standard Errors

The limiting variance-covariance matrices in the previous Propositions 2, 3 and 4, $\boldsymbol{V}_1$, $\boldsymbol{V}_2$, and $\boldsymbol{V}_3$, are difficult to estimate, as they require the computation of several numerical derivatives. Moreover, $\boldsymbol{V}_3$ reflects the contribution of parameter estimation error. As a viable route to cope with this lack of a closed-form expression for the standard errors, we develop bootstrap standard errors consistent for $\boldsymbol{V}_1$, $\boldsymbol{V}_2$, and $\boldsymbol{V}_3$. Note that our estimation procedure is an hybrid between Indirect Inference and Simulated GMM, which leads to technical issues, arising because the auxiliary models we use are potentially dynamically misspecified, and likely have a score that is not necessarily a martingale difference sequence. A natural solution is to appeal to the "block-bootstrap," described succinctly in Appendix B.2. The block-bootstrap method takes into account a possible correlation of the score of the auxiliary models. In Appendix B.2, we develop three results (Propositions B1, B2 and B3), which 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, as published by the Chicago Board of Exchange. Data for the VIX index are available daily, but only for the period following January 1990. Our macroeconomic variables include the consumer price index (CPI) and the index of industrial production (IP) for the US. Information related to the CPI and the IP index 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 index of IP, from January 1950 to December 2006, for a total of 672 observations.

We use this sample to estimate the model. 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. Such 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 events of the subprime crisis, which are quite unique in financial history and, of course, extremely relevant to the purposes of our paper; third, both realized stock market volatility and the VIX index reached record highs, as discussed below, and possibly pose serious challenges to rational models of asset prices. As we shall explain, 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 the macroeconomic data (the CPI and the index of IP) available from January 2007, and compare the predictions of the model with the actual movements in the market, stock market volatility and the VIX index.

To create the two macroeconomic factors from the CPI and the index of IP, we compute gross inflation and gross industrial production growth, both at a yearly level,

$$y_{1,t} \equiv \mathrm{CPI}_t/\mathrm{CPI}_{t-12} \quad \text{and} \quad y_{2,t} \equiv \mathrm{IP}_t/\mathrm{IP}_{t-12},$$

where CPI_t is the consumer price index and IP_t is the seasonally adjusted industrial production index, as of month t. As we explain in the Introduction, many theoretical explanations of asset price movements, and in fact, the empirical evidence, would lead us to expect that asset prices are indeed related to variables tracking the business cycle conditions (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 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 $y_{1,t}$ and $y_{2,t}$, 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 asset pricing implications of this property are then promising: although inflation has become less volatile, its persistence makes it a candidate for being a risk for the long run. The inclusion of inflation as a determinant of the pricing kernel displays one additional 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 gross 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 gross inflation and gross industrial production 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 for the joint process of the two macroeconomic variables, $y_{1,t}$ and $y_{2,t}$. The estimates are obtained through the first step of the procedure set forth in Section 3.1. In parenthesis, we report the standard errors computed through the block-bootstrap procedure developed in Appendix B.2. These estimates, which are all largely significant, confirm our previous 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 $\bar{\kappa}_2$. Moreover, the sign and value of these parameters are those we need to match the impulse-response functions for $y_{1,t}$ and $y_{2,t}$ in the data (not reported here, for space reasons). These feedback effects do have asset pricing implications, as we shall discuss below. Finally, note that the estimates of β_1 and β_2 are both negative, implying that the volatility of these two macroeconomic variables are countercyclical, another interesting property, from an asset pricing perspective.

4.2.2 Aggregate stock returns and volatility

Table 2 reports parameter estimates and standard errors for (i) the parameters linking the two macroeconomic factors, $y_{1,t}$ and $y_{2,t}$, and the unobservable factor, $y_{3,t}$, to the real stock price index, S_t; and (ii) the parameters for the unobservable factor process. Parameter estimates are obtained through the second step of our estimation strategy, explained in Section 3.2. Standard errors are computed through the block-bootstrap procedure developed in Appendix B.2. The parameter estimates are all largely significant. They point to two main conclusions. First, the stock price is positively related to both inflation and IP growth, although the link with IP growth seems to be of paramount importance. Second, the unobservable factor is largely persistent, and displays a large 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) 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.

Finally, the conditional volatility of the third factor, evaluated at the long run mean, $\mu_3 \equiv 1$, i.e. $\sqrt{\alpha_3 + \beta_3 y_3}\big|_{y_3=1}$, is about 35. This figure implies a price impact of a volatility shock in the third factor equal to $s_3 \cdot \sqrt{\alpha_3 + \beta_3 y_3}\big|_{y_3=1} = 0.38$, at the estimated parameters. In comparison, these price impacts equal $s_1 \cdot \sqrt{\alpha_1 + \beta_1 y_1}\big|_{y_1=\mu_1} = 1.48 \times 10^{-3}$, for inflation, and $s_2 \cdot \sqrt{\alpha_2 + \beta_2 y_2}\big|_{y_2=\mu_2} = 0.11$, for industrial production growth, at the estimated parameters. Hence, in our model, the price impact of a shock in the volatility of the IP growth, is about three times smaller than that resulting from the volatility of the unobservable factor. We shall discuss below the role inflation plays in the context of our estimated model.

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. For each simulation i, the stock price is computed through Eq. (28), using all the estimated parameters, the sample paths of the two macroeconomic factors, $y_{1,t}$ and $y_{2,t}$, and the simulated path of the third factor. Given the simulated stock price, we compute stock returns and volatility, for each simulation. Finally, 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. (13).

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%, with a standard deviation of about 3.8%. The model predicts an average volatility of about 13%, with a standard deviation of about 1.3%.

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. Naturally, the cyclical properties of stock returns and volatility that we see in Figure 2 can only arise due to the fluctuation of the macroeconomic factors. To quantitatively assess these properties, we perform the following experiment. We freeze the path of each factor $y_{j,t}$ at its estimated long run mean,

μ_j, simulate the model, and compute the average stock market returns, stock volatility and the standard deviation of stock volatility, over the simulations. Naturally, in this exercise, we cannot compute volatility through Eq. (13), which is based on the assumptions that the three factors are indeed stochastic, and affect the asset price. Instead, we calculate the stock volatility of the model, through the same formulae that we use to compute stock volatility from the data.

Figures 3 and 4, and the following table, report the results of this numerical experiment. When we shut down the gross inflation channel, we do not achieve any noticeable percentage reduction in the model-implied average volatility and its standard deviation. Therefore, gross inflation seems to play a marginal role as a determinant of stock market volatility, and its cyclical properties. At the same time, the 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 -42%, while the correlation our model generates is -22%.[6] Finally, the correlation between stock volatility and inflation is 23% in the data, while that implied by the model is 24%.

Percentage reduction in stock volatility
and the volatility of stock volatility

	volatility	vol of vol
without y_1	≈ 0	≈ 0
without y_2	11%	77%
without y_3	69%	-95%

Instead, industrial production growth plays a quite important role. Fixing $y_{2,t}$ at its long run mean leads to about a 10% reduction in the average level of volatility, although the third unobservable factor is key in explaining the *level* of stock volatility: when we fix $y_{3,t} = \mu_3$, thus taking the variability of $y_{3,t}$ out of the picture, the average level of volatility drops by nearly 70%. At the same time, industrial production is needed to explain the cyclical swings of stock volatility that we have in the data. When $y_{2,t}$ is frozen at μ_2, the standard deviation of the model-implied stock volatility drops dramatically by 77%.

When, instead, $y_{3,t}$ is frozen at μ_3, we even observe an *increase* in the variability of stock volatility, of about 95%. This finding is easily explained. As shown in Figure 1, gross industrial production was very volatile during the 1950s, which translates into a similar property for the aggregate market returns. Indeed, Figure 3 shows that when $y_{3,t}$ is frozen at μ_3, the model predicts

[6] In our model, the real stock price is positively related to both inflation and growth, but the correlation between inflation and growth is about -24%, which explains the negative figure our model predicts for the correlation between real returns and inflation.

that the level of stock volatility is quite high until the recession occurred in 1960, although it then progressively lowers. It is this change in level occurring during the 1960s, which makes the standard deviation of stock volatility even higher than when the unobservable factor is not frozen (as in Figure 2). If we condition on subsamples that only include the Great Moderation era (e.g., from January 1985), we find that the standard deviation of stock volatility is back to approximately 1.3%. In other words, the presence of an unobservable factor has virtually no effect on the variability of stock volatility, during the Great Moderation.

In fact, the main challenge of the model is to explain why we have observed a sustained stock market volatility, in spite of the Great Moderation. Our estimation results lead to a quite clear conclusion: the level of stock volatility cannot be explained by macroeconomic variables only. Instead, some unobservable factor is needed, which accounts for about two thirds of the fluctuations in stock volatility (precisely, 69%). At the same time, the same unobservable factor cannot explain the variability in stock volatility. As Figure 4 reveals, when we freeze the two macroeconomic factors at their unconditional means, μ_1 and μ_2, the stock volatility predicted by the model is just the average of its empirical counterpart, with virtually no fluctuations at all.[7] All in all, our empirical results suggest that the volatility of stock volatility is attributable to the *cyclical* variations in stock volatility, which our model captures through the relation between asset returns and industrial production growth. Of course, these swings are amplified by the presence of the third unobservable factor.

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

Table 3 reports parameter estimates and standard errors for the vector of the risk-premiums coefficients $\boldsymbol{\lambda}$ in the risk-premium process of Eq. (19). The estimates, which are all significant, are obtained through the third, and final step, of the estimation procedure, described in Section 3.3. Standard errors are computed through the block-bootstrap procedure developed in Appendix B.2.

The estimates imply that the risk-premiums processes are all positive, and quite large. Moreover, the risk compensation for gross 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. (19). While gross inflation does receive compensation, and helps explain the dynamics of the volatility risk-premium, as we explain below, the countercyclical behavior of the risk-premium for industrial production growth is even more critical, at least over the period

[7] In Figure 4, stock volatility fluctuates between around 11% and 12%, and yet stock returns are quite smooth. The reason is that stock returns are computed yearly, while stock volatility is estimated with one-month returns, which the model predicts to be quite volatile (precisely, about 12%, annualized, on average), even in the absence of macroeconomic factors.

from 1990 to 2006. Our estimated model predicts indeed 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 reproduce well the large swings in the VIX index that we have observed during the 1991 and the 2001 recession episodes.

The top panel of Figure 5 also shows the dynamics of expected future volatility, 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 the risk-premiums required to bear the fluctuations of the macroeconomic factors are (i) positive and (ii) countercyclical, as argued above, and, also, because (iii) current volatility is countercyclical. Expected future volatility is countercyclical, under the physical probability, only because of the third effect. Figure 6 reveals the "tilt" in the future paths of industrial production growth that we need, in order to make our model match the data on the VIX index. The left panel of this picture depicts sample paths over one month, under the physical probability. The right panel depicts sample paths under the risk-neutral probability. In words, movements in the volatility risk-premiums account for the variation of the VIX index sensibly more than movements in the future expected volatility under the physical probability, as clearly summarized by Figure 5.

The substantial wedge between expected volatility under the two probabilities is actually reinforced by the feedback between inflation and industrial production growth. The mechanism is the following. Gross inflation requires a large and positive risk-premium, as already mentioned. This premium is actually so large, to make expected inflation always decrease from its current levels, under the risk-neutral probability. (Technically, the inflation risk-premium is such that the drift of inflation is negative under the risk-neutral probability.) Since the feedback parameter, $\bar{\kappa}_2$, is positive and large in value, the risk-neutral expectation of industrial production worsens even more than in the absence of feedback effects, due precisely to the inflation risk-premium. In other words, inflation affects the VIX index through a subtle channel, the compensation for the risk of correlation between inflation and growth, arising due to sizeable values of (i) the feedback between inflation and growth, as summarized by $\bar{\kappa}_2$, (ii) the inflation risk-premium. 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: while our previous findings suggest that inflation does not affect too much the dynamics of stock returns and volatility, the presence of significant feedbacks between inflation and industrial production growth, in conjunction with compensation for inflation risk, reveal that inflation does affect future expected volatility, under the risk-neutral probability.

Finally, the bottom panel in Figure 5 plots the volatility risk-premium, defined as the difference between the (square roots of the) model-implied expected integrated variance under the risk-neutral probability and the model-implied expected integrated variance under the physical probability. This risk-premium is countercyclical, and this property arises for exactly the same reasons we put forward to explain the large swings of the VIX index predicted by the model: positive compensation for risk, countercyclical variation of the risk-premiums required to compensate for the risk in fluctuations of the macroeconomic factors, and feedback effects between the two macroeconomic factors. Interestingly, the two recessions, in 1991 and 2001, seem to be anticipated by a surge in the volatility risk-premium. Figure 7 provides scatterplots of the volatility risk-premium against inflation and industrial production. The top panel reveals that the volatility risk-premium does not display a neat relation with inflation, although we have explained that the inflation risk-premium is an important determinant of it, as it magnifies the industrial production growth channel, through a feedback effect. Instead, the bottom panel reveals a neat and inverse relation between the volatility risk-premium and industrial production growth, and suggests an interesting "convexity" property: in good times, the volatility risk-premium does not move too much in response of movements in the industrial production growth, although it increases quickly and by a sizeable amount as soon as business cycle conditions deteriorate.

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 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. (29), 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 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. The shaded area in Figure 1 covers the out-of-sample behavior of gross inflation and growth.

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. We now feed the model with the macroeconomic factors observed from January 2007 to March 2009, to obtain quantitative predictions about stock returns, stock volatility and the VIX index.

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 -23.28%, which is the second lowest figure our model produces, overall. (The lowest level the model predicts is -23.64%, for June 1954, when both inflation and industrial production growth were quite low.) Instead, the model predicts that stock volatility and the VIX index surge even more than in the data, reaching record highs of 32.48% (volatility) and 73.67% (VIX).

Figure 8 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. (31), excluding the lag for six months, related to the autoregressive term. The OLS for the VIX index is that in Eq. (34). 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 for the out-of-sample period.

RMSE for the model and OLS

	Model	OLS
Volatility	0.0508	0.0700
VIX Index	0.1215	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. (34), an autoregressive, distributed lag model) produces predictions that are not as accurate as the model, and generate overfit. The model, instead, reproduces the huge swings we see in the data, in both the last two recession episodes. It also seems to anticipate turning points, in that it raises before and drops at the end of a recession. The RMSEs clearly favour the model against OLS, although it appears to do so more with volatility than for the VIX, as Figure 8 informally reveals.

5 Conclusion

How precisely 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 shows that approximately one third of this level can be explained by macroeconomic factors, and that some unobserved component is needed indeed to make stock volatility consistent with rational asset valuation. Moreover, we show that a business cycle factor is needed to explain the inevitable fluctuations of stock volatility around its average.

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 in the macroeconomic factors is large and countercyclical, and explains the large 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 successfully reproduces, through a countercyclical variation in the volatility risk-premiums. Finally, we provide evidence that the same volatility risk-premiums might help predict developments in the business cycle in bad times—the end of a recession.

Which macroeconomic factor matters? We find that industrial production growth is largely responsible for the random fluctuations of stock volatility around its level, and that inflation plays, instead, a quite limited role in this context. At the same time, inflation plays an important role as a determinant of the VIX index, through two channels: (i) one, direct, channel, related to the inflation risk-premium, and (ii) an indirect channel, arising from the business cycle propagation mechanism, through which inflation and industrial production growth are correlated. The second channel is subtle, as it gives rise to a correlation risk that we show it is significantly priced by the market.

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. An even more 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 J. Liu, 2004. "How to Discount Cashflows with Time-Varying Expected Returns," *Journal of Finance* 59, 2745-2783.

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.

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

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.

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

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.

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.

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, 123-156.

Bollerslev, T., M. Gibson and H. Zhou, 2004. "Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option Implied and Realized Volatilities," Federal Reserve Board WP 2004-56.

Bollerslev, T., G. Tauchen and H. Zhou, 2008. "Expected Stock Returns and Variance Risk Premia," forthcoming in the *Review of Financial Studies.*

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.

Drechsler, I. and A. Yaron, 2008. “What’s Vol Got to Do With It,” working paper Wharton School, University of Pennsylvania.

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. 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.

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, 2007. “Estimation of Objective and Risk-Neutral Distributions Based on Moments of Integrated Volatility,” forthcoming in the *Journal of Econometrics.*

Glasserman, P. and K. K. Kim, 2008. “Moment Explosion and Stationary Distributions in Affine Diffusion Models,” forthcoming in *Mathematical Finance.*

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

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

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

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.

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

Mamaysky, H., 2002. “A Model for Pricing Stocks and Bonds,” working paper Yale School of Management.

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.

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

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.

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., 2009. “Variance Risk Premium Dynamics: The Role of Jumps,” forthcoming in the *Review of Financial Studies.*

Todorov, V. and G. Tauchen, 2009. “Volatility Jumps,” working paper Duke University and Northwestern University.

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. (25), and (ii) the sample mean and standard deviation of $y_{1,t}$ and $y_{2,t}$. The block-bootstrap procedure for the standard errors of the estimates is developed in Appendix B.2. The sample covers monthly data for the period from January 1950 to December 2006.

	Estimate	Std error
κ_1	0.0331	$3.4630 \cdot 10^{-4}$
μ_1	1.0379	$3.4855 \cdot 10^{-3}$
α_1	$2.2206 \cdot 10^{-4}$	$2.7607 \cdot 10^{-6}$
β_1	$-9.6197 \cdot 10^{-7}$	$1.0099 \cdot 10^{-8}$
κ_2	0.5344	$7.4482 \cdot 10^{-3}$
μ_2	1.0415	$4.9926 \cdot 10^{-3}$
α_2	0.0540	$3.5233 \cdot 10^{-4}$
β_2	-0.0497	$3.3939 \cdot 10^{-4}$
$\bar{\kappa}_1$	-0.2992	$4.3054 \cdot 10^{-3}$
$\bar{\kappa}_2$	1.2878	$1.8091 \cdot 10^{-2}$

Table 2

Parameter estimates and block-bootstrap standard errors for the stock price and the unobservable factor:

$$s(t) = s_0 + \sum_{i=1}^{3} s_i y_i(t),$$

where $s(t)$ is the real stock price, obtained as the ratio between the S&P Compounded index and the Consumer Price Index; $y_1(t)$ and $y_2(t)$ are the observed gross inflation and gross industrial production growth, as defined in Table 1; finally, $y_3(t)$ is an unobserved factor, with the following dynamics:

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

where $W_3(t)$ is a standard Brownian motion, and is independent of the Brownian motions driving the fluctuations of the two macroeconomic factors $y_1(t)$ and $y_2(t)$. The parameter vector is $\boldsymbol{\theta}^\top = (\kappa_3, \mu_3, \alpha_3, \beta_3, s_0, s_j, j = 1, 2, 3)$, 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. Matching conditions relate to (i) parameter estimates for the auxiliary model for stock returns, Eq. (30), and for the auxiliary model for stock volatility, Eq. (31), and (ii) the sample mean and standard deviation of stock returns and volatility. The block-bootstrap procedure for the standard errors of the estimates is developed in Appendix B.2. The sample covers monthly data for the period from January 1950 to December 2006.

	Estimate	Std error
s_0	0.1279	$6.3962\cdot10^{-2}$
s_1	0.0998	$4.9570\cdot10^{-2}$
s_2	2.5103	$1.5668\cdot10^{-1}$
s_3	0.0109	$4.6518\cdot10^{-3}$
κ_3	0.0092	$2.8930\cdot10^{-3}$
μ_3	1	restricted
α_3	$9.4543\cdot10^{2}$	$3.2877\cdot10^{2}$
β_3	4.1653	2.0533

Table 3

Parameter estimates and block-bootstrap standard errors for the risk-premium parameters of the total risk-premium process in Eq. (19):

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

where $y_1(t)$ and $y_2(t)$ are the observed 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. (34), and (ii) the sample mean and standard deviation of the VIX index. The block-bootstrap procedure for the standard errors of the estimates is developed in Appendix B.2. The sample covers monthly data for the period from January 1990 to December 2006.

	Estimate	Std error
Inflation		
$\lambda_{1(1)}$	$-24.6491\cdot10^2$	$31.4150\cdot10$
$\lambda_{2(1)}$	36.5596	4.8917
Ind. Prod.		
$\lambda_{1(2)}$	$47.0883\cdot10$	24.0197
$\lambda_{2(2)}$	1.1159	0.1677
Unobs.		
$\lambda_{1(3)}$	-0.2078	0.0274
$\lambda_{2(3)}$	$27.7137\cdot10$	35.6533

Figures

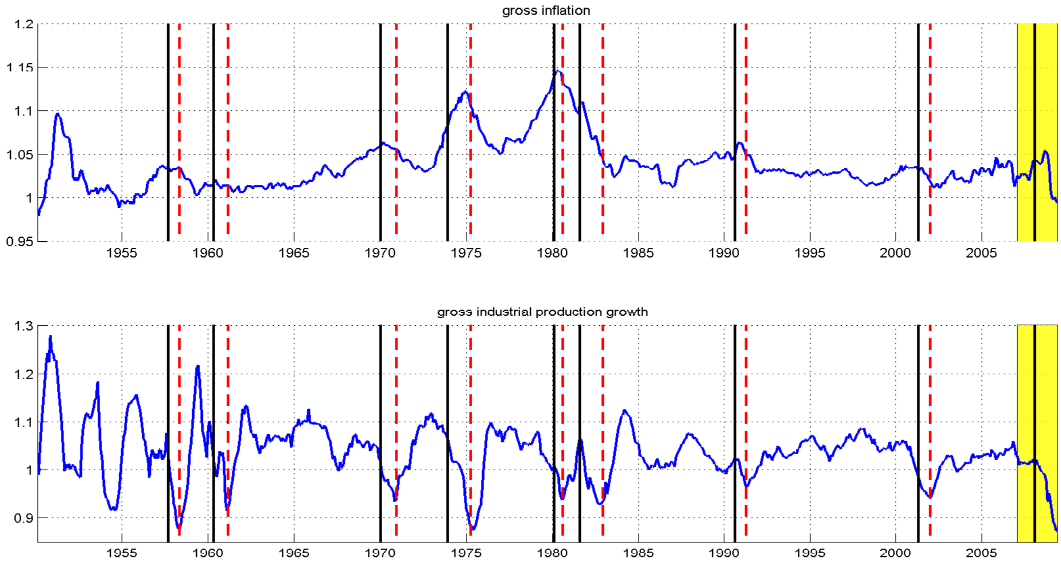


Figure 1 – Industrial production growth and inflation, with NBER dated recession periods. This figure plots the one-year, monthly gross inflation, defined as $y_{1,t} \equiv \mathrm{CPI}_t / \mathrm{CPI}_{t-12}$, and the one-year, monthly gross industrial production growth, defined as $y_{2,t} \equiv \mathrm{IP}_t / \mathrm{IP}_{t-12}$, where CPI_t is the Consumer price index as of month t, and IP_t is the real, seasonally adjusted industrial production index as of month t. The sample covers monthly data for the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions. The shaded area (in yellow) covers the out-of-sample period, from January 2007 to March 2009, which we use to formulate model's predictions.

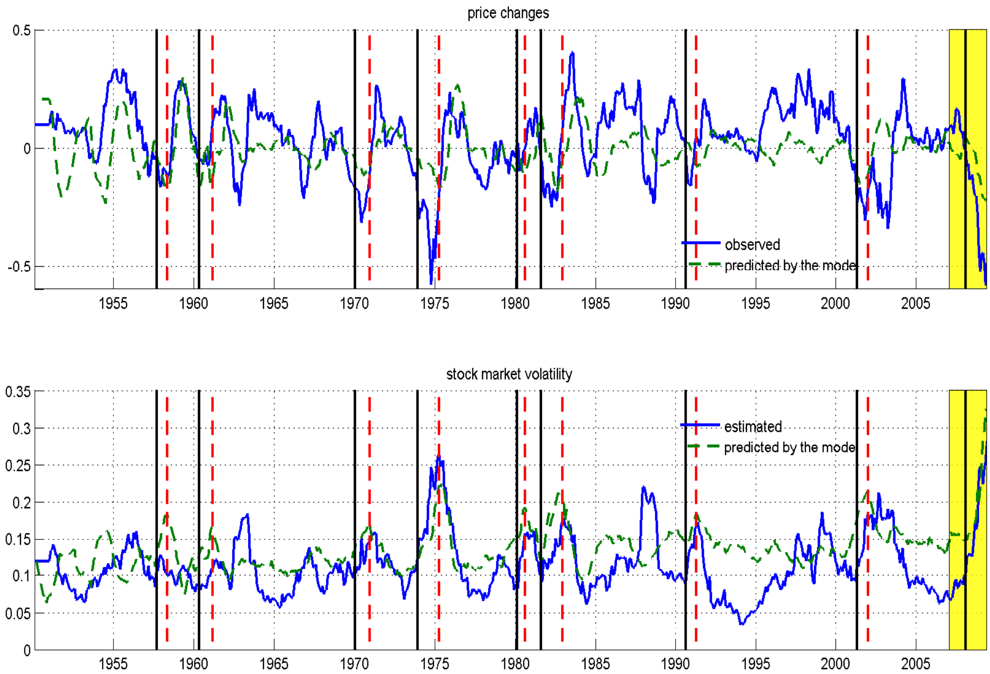
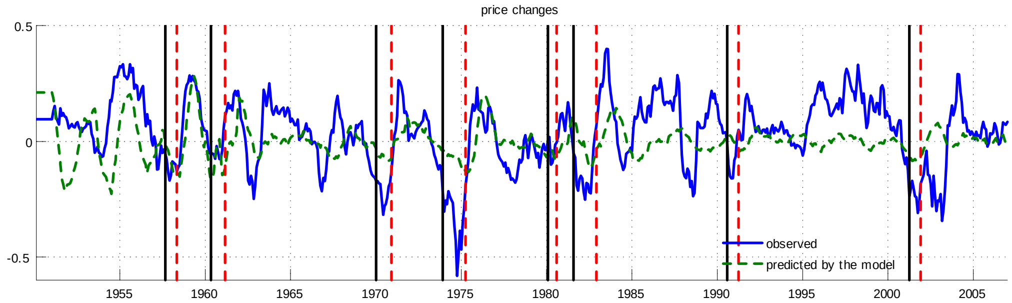


Figure 2 – Stock returns and volatility along with the model predictions, with NBER dated recession periods, and out-of-sample predictions. This figure plots one-year ex-post price changes and one-year return volatility, along with their counterparts predicted by the model. The top panel depicts continuously compounded price changes, defined as $R_t \equiv \ln(s_t/ s_{t-12})$, where s_t is the real stock price as of month t. The bottom panel depicts smoothed return volatility, defined as $\mathrm{Vol}_t \equiv \sqrt{6\pi} \cdot 12^{-1} \sum_{i=1}^{12} |\ln(s_{t+1-i}/s_{t-i})|$, along with the instantaneous standard deviation predicted by the model, obtained through Eq. (12). Each prediction at each point in time is obtained by feeding the model with the two macroeconomic factors depicted in Figure 1 (inflation and growth) 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. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions. The shaded area (in yellow) covers the out-of-sample period, from January 2007 to March 2009, which we use to formulate model's predictions.



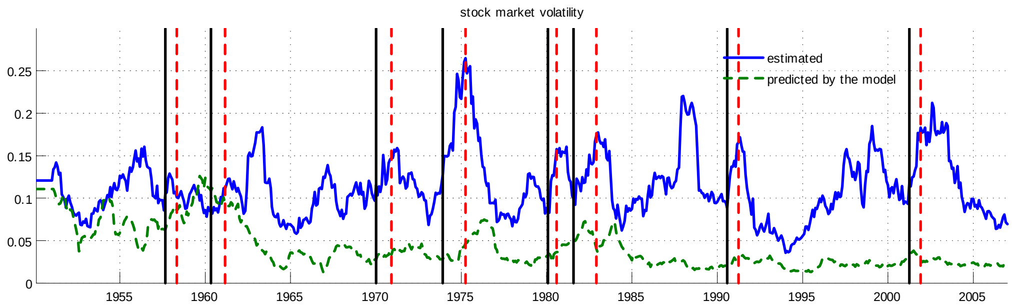


Figure 3 – Stock returns and volatility along with the model predictions in the absence of an unobservable factor, with NBER dated recession periods. This figure plots one-year price changes and one-year return volatility, along with their counterparts predicited by the model, when the model is driven by the macroeconomic factors only. The top panel depicts continuously compounded price changes, defined as $R_t \equiv \ln(s_t/s_{t-12})$, where s_t is the real stock price as of month t. The bottom panel depicts smoothed return volatility, defined as $\mathrm{Vol}_t \equiv \sqrt{6\pi} \cdot 12^{-1} \sum_{i=1}^{12} |\ln(s_{t+1-i}/s_{t-i})|$, along with the instantaneous standard deviation predicted by the model, estimated in the same way as for the data. Each prediction at each point in time is obtained by feeding the model with the two macroeconomic factors depicted in Figure 1 (inflation and growth) and by freezing the unobserved factor at its long run mean, $\mu_3 = 1$. The sample covers monthly data for the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions.

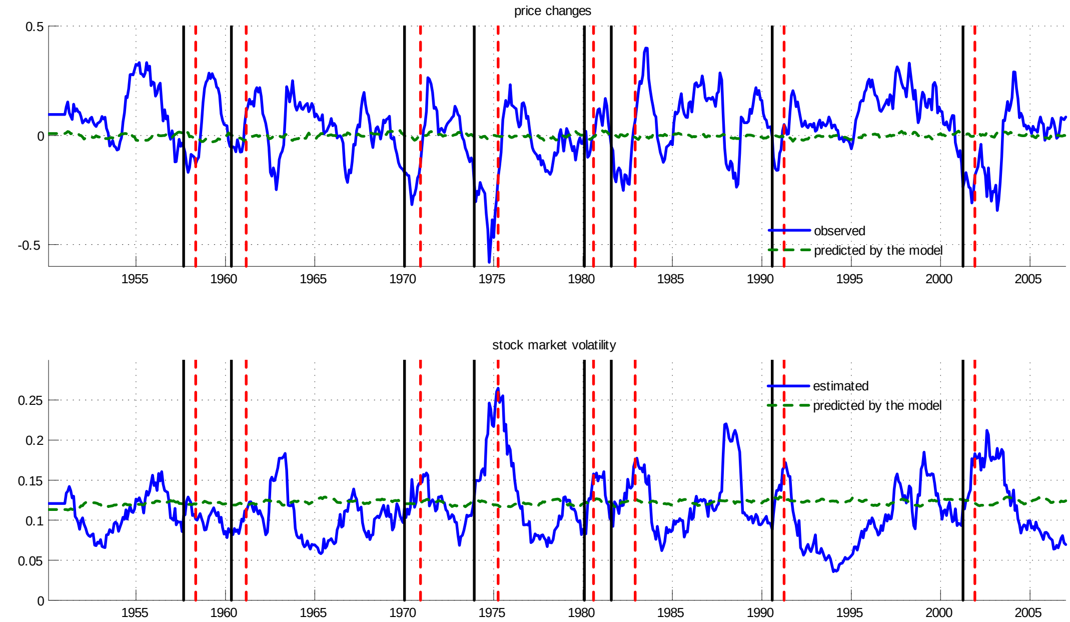


Figure 4 – Stock returns and volatility along with the model predictions in the absence of macroeconomic factors, with NBER dated recession periods. This figure plots one-year price changes and one-year return volatility, along with their counterparts predicited by the model, when the model is driven by the unobservable factor only. The top panel depicts continuously compounded price changes, defined as $R_t \equiv \ln (s_t / s_{t-12})$, where s_t is the real stock price as of month t. The bottom panel depicts smoothed return volatility, defined as $\mathrm{Vol}_t \equiv \sqrt{6\pi} \cdot 12^{-1} \sum_{i=1}^{12} |\ln (s_{t+1-i}/s_{t-i})|$, along with the instantaneous standard deviation predicted by the model, estimated in the same way as for the data. Each prediction at each point in time is obtained by freezing the two macroeconomic factors at their long run means, μ_1 and μ_2, and by averaging over the cross-section 1000 dynamic simulations of the unobserved factor. The sample covers monthly data for the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions.

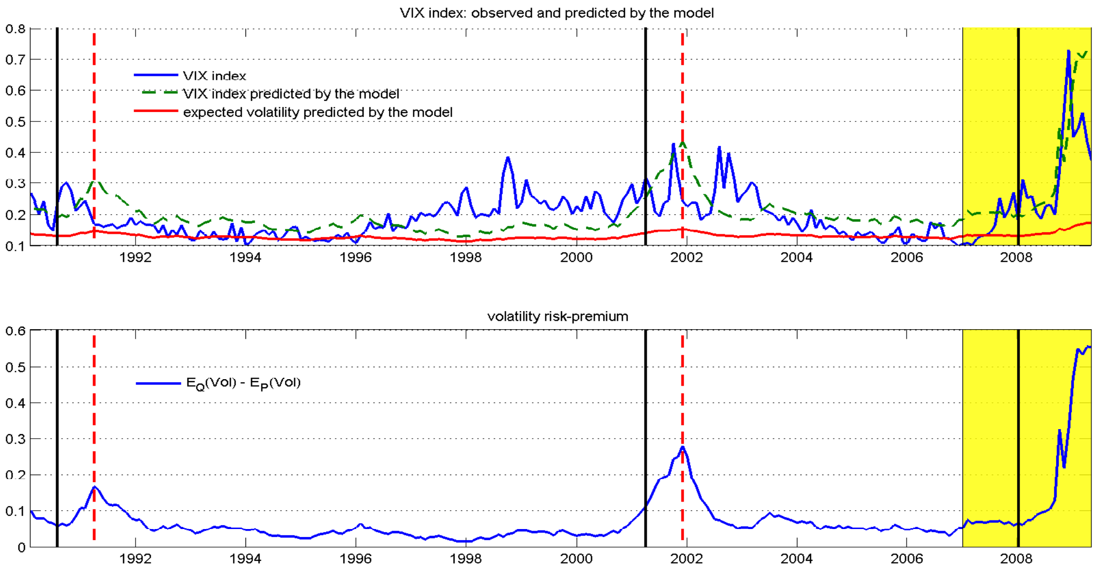


Figure 5 – The VIX Index and volatility risk-premia, with NBER dated recession periods, and out-of-sample predictions. This figure plots the VIX index along with model's predictions. The top panel depicts (i) the VIX index, (ii) the VIX index predicted by the model, and (iii) the VIX index predicted by the model in an economy without risk-aversion, i.e. the expected integrated volatility under the physical probability. The bottom panel depicts the volatility risk-premium predicted by the model, defined as the difference between the model-generated expected integrated volatility under the risk-neutral and the physical probability,

$$\mathrm{VRP}\left(\boldsymbol{y}\left(t\right)\right) \equiv \sqrt{\frac{1}{T-t}}\left(\sqrt{\mathbb{E}\left(\int_t^T \sigma^2\left(\boldsymbol{y}\left(u\right)\right)\mathrm{d}u\,\middle|\,\boldsymbol{y}\left(t\right)\right)}-\sqrt{E\left(\int_t^T \sigma^2\left(\boldsymbol{y}\left(u\right)\right)\mathrm{d}u\,\middle|\,\boldsymbol{y}\left(t\right)\right)}\right),$$

where $T-t=12^{-1}$, $\mathbb{E}$ is the conditional expectation under the risk-neutral probability, E is the conditional expectation under the true probability, $\sigma^2(\boldsymbol{y})$ is the instantaneous variance predicted by the model, obtained through Eq. (12), and $\boldsymbol{y}$ is the vector of three factors: the two macroeconomic factors depicted in Figure 1 (inflation and growth) and one unobservable factor. Each prediction at each point in time is obtained by feeding the model with the two macroeconomic factors depicted in Figure 1 (inflation and growth) 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 1990 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions. The shaded area (in yellow) covers the out-of-sample period, from January 2007 to March 2009, which we use to formulate model's predictions.

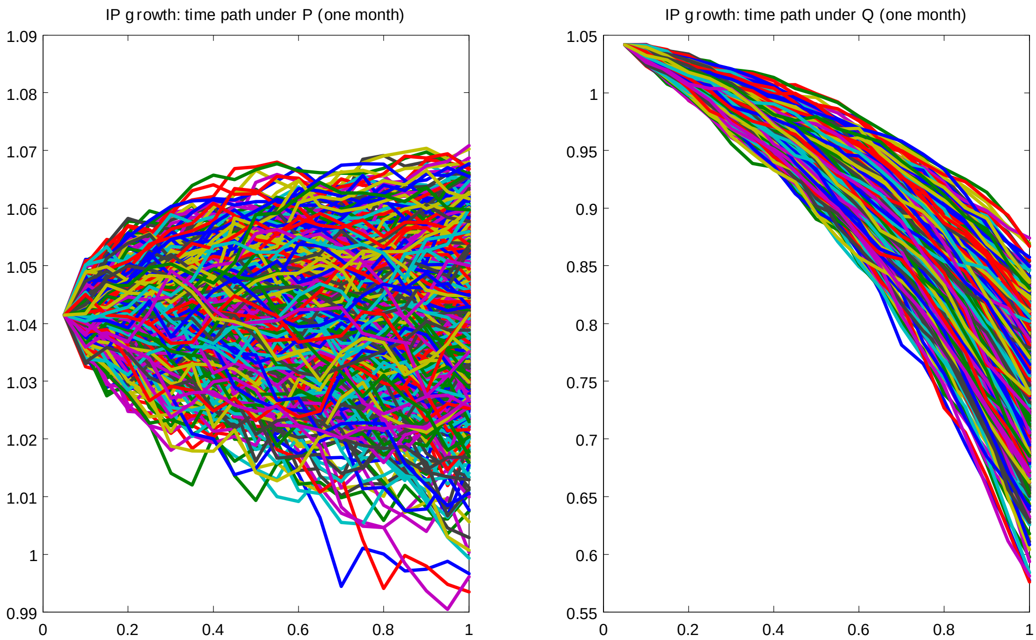


Figure 6 – Sample paths of industrial production growth: under the objective probability, and implied by the VIX index. This figure plots 1000 simulations of one month paths of the gross industral production growth, with starting values fixed at 1.03 (gross inflation) and 0.96 (gross industrial production growth). The left panel displays the sample paths under P, the physical probability. The right panel depicts the sample paths under Q, the risk-neutral probability, obtained by matching the model to the VIX index.

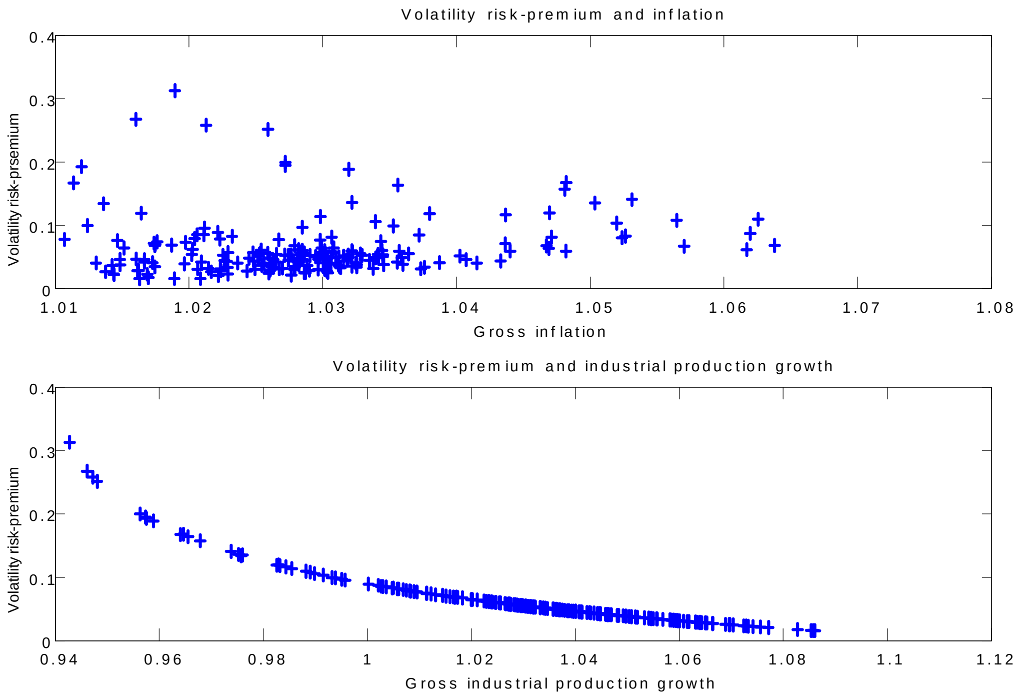


Figure 7 – Volatility risk-premium against inflation and industrial production growth. This figure provides scatterplots of the volatility risk-premium predicted by the model, depicted in Figure 3 (bottom panel), against the two macroeconomic factors depicted in Figure 1 (inflation and growth). Each prediction at each point in time is obtained by feeding the model with the two macroeconomic and by averaging over 1000 dynamic simulations of the unobserved factor. The sample covers monthly data for the period from January 1990 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions.

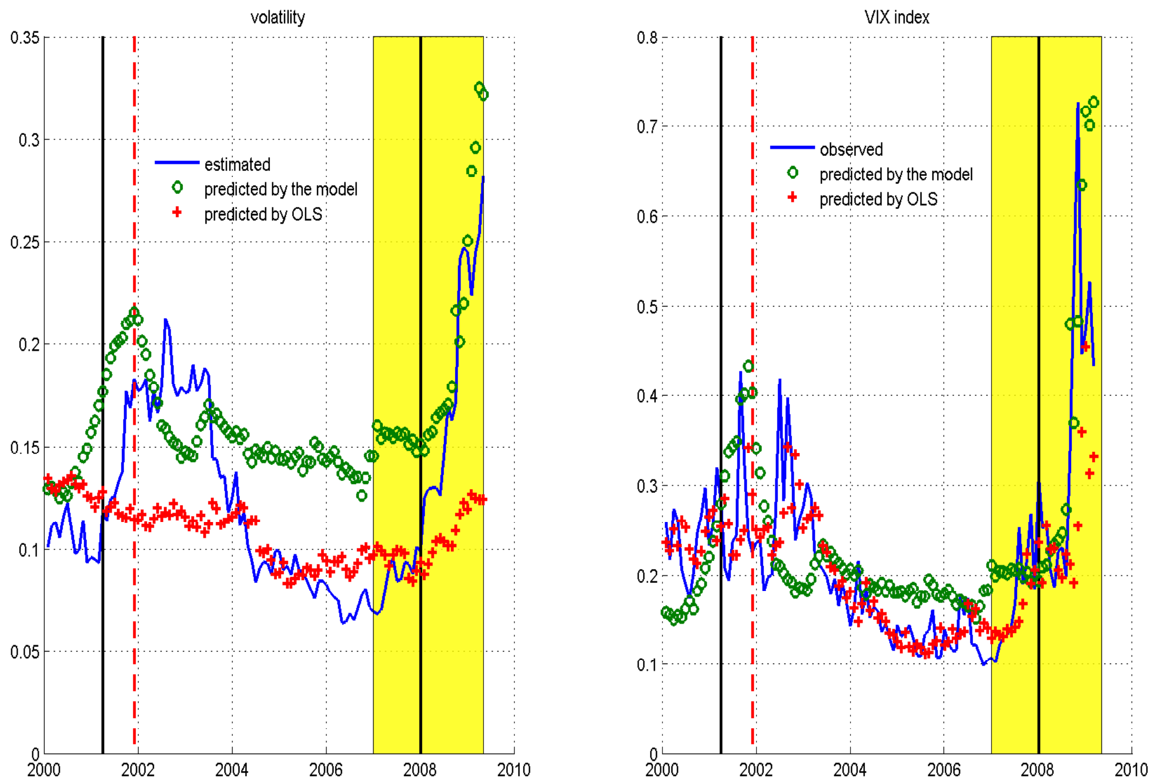


Figure 8 – Out of sample predictions and the subprime crisis. This figure plots one-year return volatility and the VIX index, along with its counterparts predicted by the model and by an OLS regression. The left panel depicts smoothed return volatility, defined as $\text{Vol}_t \equiv \sqrt{6\pi} \cdot 12^{-1} \sum_{i=1}^{12} |\ln (s_{t+1-i}/s_{t-i})|$, where s_t is the real stock price as of month t, along with the instantaneous standard deviation predicted by (i) the model, through Eq. (12), and (ii) the predictive part of an OLS regression of Vol_t on to past values of Vol_t, inflation and industrial production growth. The right panel depicts the VIX index, along with the VIX index predicted by (i) the model; and (ii) the predictive part of an OLS regression of the VIX index on to past values of the VIX index, inflation and industrial production growth. Each prediction is obtained by feeding the model and the predictive part of the OLS regression with the two macroeconomic factors depicted in Figure 1 (inflation and growth) and, for the model, by averaging over the cross-section of 1000 dynamic simulations of the unobserved factor. The sample depicted in the figure spans the period from January 2000 to March 2009. The estimation of both the model and the OLS regressions relates to the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and the vertical dashed line (in red) indicates the end of the NBER-dated recession, occurred in November 2001. The shaded area (in yellow) covers the out-of-sample period, from January 2007 to March 2009, which includes the NBER recession announced to have occurred in December 2007, and the subprime crisis, which started in June 2007.

Technical Appendix

A. Proofs for Section 2

Existence of a strong solution to Eq. (1) and Eq. (18).

Consider the following conditions: for all i,

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

(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. (1) for which $V(\boldsymbol{y}(t))_{(ii)} > 0$ for all t almost surely.

We apply these conditions to the diffusion in Eq. (18). 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 $\beta_i = 0$,

$$\kappa_i (\mu_i \beta_i + \alpha_i) + \bar{\kappa}_i \beta_i \left(\mu_j + \frac{\alpha_j}{\beta_j}\right) > \frac{1}{2}\beta_i^2, \quad i \neq j. \tag{A1}$$

Proof of Proposition 1

The technical condition in Proposition 1 is,

$$E\left[\int_t^T \left\| \frac{\boldsymbol{\eta}^\top \boldsymbol{\Sigma} \boldsymbol{V}(\boldsymbol{y}(\tau))}{\gamma + \boldsymbol{\eta}^\top \boldsymbol{y}(\tau)} - \boldsymbol{\Lambda}(\tau)^\top \right\|^2 \mathrm{d}\tau\right] < \infty, \tag{A2}$$

for some constants γ and $\boldsymbol{\eta}$ in Eq. (A10) below.

Next, 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,

$$\frac{\mathrm{d}s^\xi(t)}{s^\xi(t)} = \mathrm{Dr}(\boldsymbol{y}(t))\,\mathrm{d}t + \left(\boldsymbol{Q}(\boldsymbol{y}(t))^\top - \boldsymbol{\Lambda}(\boldsymbol{y}(t))^\top\right)\mathrm{d}\boldsymbol{W}(t), \tag{A3}$$

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 $T < \infty$,

$$s^\xi(t) = E\left[\int_t^T \delta^\xi(h)\,\mathrm{d}h \middle| \mathbb{F}(t)\right] + E[s^\xi(T) \mid \mathbb{F}(t)], \tag{A4}$$

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,

$$\lim_{T\to\infty} E[s^\xi(T) \mid \mathbb{F}(t)] = 0, \tag{A5}$$

from which Eq. (8) in the main text follows, once we show that $\int_t^\infty E[\delta^\xi(h)]\mathrm{d}h < \infty$.

Next, by Eq. (A4),

$$0 = \frac{\mathrm{d}}{\mathrm{d}\tau} E[s^\xi(\tau) \mid \mathbb{F}(t)]\bigg|_{\tau=t} + \delta^\xi(t). \tag{A6}$$

Below, we show that

$$E[s^\xi(T) \mid \mathbb{F}(t)] = s^\xi(t) + \int_t^T \mathrm{D}(\boldsymbol{y}(h))\, s^\xi(h)\,\mathrm{d}h. \tag{A7}$$

Therefore, by the assumptions on $\mathbf{\Lambda}$, Eq. (A6) can be rearranged to yield the following ordinary differential equation,

$$\text{For all } \boldsymbol{y}, \quad s_y(\boldsymbol{y})^\top (\boldsymbol{c} - \boldsymbol{D}\boldsymbol{y}) + \frac{1}{2}\mathrm{Tr}\left([\mathbf{\Sigma}\boldsymbol{V}(\boldsymbol{y})][\mathbf{\Sigma}\boldsymbol{V}(\boldsymbol{y})]^\top s_{yy}(\boldsymbol{y})\right) + \delta(\boldsymbol{y}) - rs(\boldsymbol{y}) = 0, \tag{A8}$$

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

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

$$s(\boldsymbol{y}) = \gamma + \boldsymbol{\eta}^\top \boldsymbol{y}, \tag{A9}$$

for some scalar γ and some vector $\boldsymbol{\eta}$. By plugging this guess back into Eq. (A8) 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,

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

The solution to this system is,

$$\gamma = \frac{\delta_0 + \boldsymbol{\eta}^\top \boldsymbol{c}}{r} \quad \text{and} \quad \boldsymbol{\eta}^\top = \boldsymbol{\delta}^\top (\boldsymbol{D} + r\boldsymbol{I}_{n\times n})^{-1}. \tag{A10}$$

We are left to show that Eq. (A5) and (A7) hold true.

As for Eq. (A5), 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. (A9), and the third line holds because $E[\xi(T) \mid \mathbb{F}(t)] = 1$ for all T, and by a change of measure. Eq. (A5) follows because $\boldsymbol{y}$ is stationary mean-reverting under the risk-neutral probability.

To show that Eq. (A7) holds, we need to show that the diffusion part of s^ξ in Eq. (A3) is a martingale, not only a local martingale, which it does whenever for all T,

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

which is the condition in (A2).

B. Proofs for Section 3

Remarks on notation: Hereafter, we let Avar and Acov denote the probability 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. Proofs of Propositions 2, 3 and 4

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

$$\mathbf{\Phi} = \{\boldsymbol{\phi} : \text{The inequality in (A1) holds, } \kappa_i > 0, \text{ and } \kappa_i\kappa_j - \bar{\kappa}_i\bar{\kappa}_j > 0,\ i,j = 1,2 \text{ and } i \neq j\},$$

and

$$\mathbf{\Theta} = \{\boldsymbol{\theta} : \text{The inequality in (A1) holds for } i = 3, \text{ and } \kappa_3 > 0\}.$$

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

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

and

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

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,

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

We are now ready to prove the propositions in Section 3. 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. (18) for $j = 1, 2$; (iii) As for Eq. (18), for $i, j = 1, 2$ $i \neq j$, $\kappa_i\kappa_j - \overline{\kappa}_i\overline{\kappa}_j > 0$ 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. (21); (v) The risk-premium vector $\boldsymbol{\pi}\left(\boldsymbol{y}\right)$ and the dividend vector $\delta\left(\boldsymbol{y}\right)$ are defined as in Eqs. (19) and (20).

Proof of Proposition 2

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 (2008), 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. (18), $(y^{\phi}_{1,t,\Delta,h}, y^{\phi}_{2,t,\Delta,h})$ is at least twice continuously differentiable in any open neighborhood of $\boldsymbol{\phi}_0$.

We claim that $\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0 = o_p(1)$, which follows by the usual arguments relying on unique identifiability (ensured by the previous properties of the diffusion in Eq. (18) and its simulated skeleton), and the uniform law of large numbers. 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

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

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

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

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

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

Let $\boldsymbol{Y}^{\phi_0}_{t,\Delta,h}$ be the vector containing all the regressors in Eq. (26), and let $\hat{\boldsymbol{\varphi}}_{1,T,\Delta,h}\left(\boldsymbol{\phi}_0\right)$ be the parameter estimator of the OLS regression of $y^{\phi_0}_{1,t,\Delta,h}$ on $\boldsymbol{Y}^{\phi_0}_{t,\Delta,h}$. We have:

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

$$= \left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_{t,h}^{\phi_0}\boldsymbol{Y}_{t,h}^{\phi_0\top}\right)^{-1}\sqrt{T}\left(\frac{1}{T}\sum_{t=25}^{T}\left(\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}y_{1,t,\Delta,h}^{\phi_0}-\boldsymbol{Y}_{t,h}^{\phi_0}y_{1,t,h}^{\phi_0}\right)\right)$$

$$+\sqrt{T}\left(\left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}\boldsymbol{Y}_{t,\Delta,h}^{\phi_0\top}\right)^{-1}-\left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_{t,h}^{\phi_0}\boldsymbol{Y}_{t,h}^{\phi_0\top}\right)^{-1}\right)\left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_{t,\Delta,h}^{\phi_0}y_{1,t,\Delta,h}^{\phi_0}\right). \tag{B1}$$

As for the first term on the RHS of (B1), $\left(\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_{t,h}^{\phi_0}\boldsymbol{Y}_{t,h}^{\phi_0\top}\right)^{-1}=O_p(1)$, 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 right hand side of Eq. (B1) can be dealt with similarly. Thus, we have:

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

where,

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

The last term of the right hand side 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,

$$\begin{aligned}&\mathrm{Avar}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)\right)\\&=\frac{1}{H^2}\sum_{h=1}^{H}\mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)\right)+\frac{1}{H^2}\sum_{h=1}^{H}\sum_{h'\neq h}^{H}\mathrm{Acov}\left(\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right),\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{T,h'}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)\right)\\&=\frac{1}{H}\mathrm{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{T,h}(\boldsymbol{\phi}_0)-\boldsymbol{\varphi}_0\right)\right),\text{ for all } h.\end{aligned}$$

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

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

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

Proof of Proposition 3

By the same arguments utilized in the proof of Proposition 2,

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

where

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

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

$$\boldsymbol{J}_0 = \text{Avar}\left(\sqrt{T}\left(\tfrac{1}{H}\sum_{h=1}^{H}\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0\right) - \boldsymbol{\vartheta}_0\right)\right) + \text{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right)$$
$$- 2\text{Acov}\left(\sqrt{T}\left(\tfrac{1}{H}\sum_{h=1}^{H}\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0\right) - \boldsymbol{\vartheta}_0\right), \sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right).$$

Let $\hat{\boldsymbol{\vartheta}}_{T,h}\left(\boldsymbol{\theta}_0\right)$ be the the unfeasible estimator, obtained by simulating continuous paths for the unobservable factor $Z\left(t\right)$. By the same arguments as those in the proof of Proposition 2,

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

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

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

where

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

and

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

Therefore, using the fact that $\text{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right) = \text{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,1}\left(\boldsymbol{\theta}_0\right) - \boldsymbol{\vartheta}_0\right)\right) = \boldsymbol{J}_2$, and exploiting the expression for $\boldsymbol{J}_0$, we obtain:

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

and, hence:

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

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 $y_{1,t}$, $y_{2,t}$, $t = T - \mathcal{T}, \cdots, \mathcal{T} - 1$, 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. (18),

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

where $\pi(y_i)$ denotes the i-th element of the vector $\boldsymbol{\pi}(\boldsymbol{y})$ in Eq. (19), 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. (21)-(24), the model-based volatility under the risk-neutral measure, at the j-th simulation, is:

$$\sigma^2_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda}) = \frac{\sum_{i=1}^{3} \hat{s}^2_{i,T} \left(\hat{\alpha}_{i,T} + \hat{\beta}_{i,T} \hat{y}^{\lambda}_{i,t+k\hat{\Delta},h}\right)}{\tilde{s}^2_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda})}, \tag{B2}$$

where

$$\tilde{s}_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda}) = \hat{s}_{0,T} + \sum_{i=1}^{3} \hat{s}_{i,T} \hat{y}^{\lambda}_{i,t+k\hat{\Delta},h}, \tag{B3}$$

and $\hat{s}_{l,T}$ $l = 0, \cdots, 3$ are the reduced-form parameters obtained in step 2 of the estimation procedure. Finally, we compute the simulated value of the model-based VIX, $\mathrm{VIX}_{t,\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda})$, by integrating volatility over each month, as follows:

$$\mathrm{VIX}_{t,\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda}) = \sqrt{\frac{1}{\hat{\Delta}} \sum_{k=0}^{\hat{\Delta}^{-1}-1} \sigma^2_{t+(k+1)\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \boldsymbol{\lambda})}. \tag{B4}$$

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 $\hat{\Delta}$.

Proof of Proposition 4

Given Assumptions B1(i) and B1(iii), for any $\boldsymbol{\lambda}$ in a compact set $\boldsymbol{\Lambda}_0$, $y^{\lambda}_{i,t+(k+1)\Delta,h}$, $i = 1, 2, 3$, $h = 1, \cdots, H$, is geometrically β-mixing, and has a stationary distribution with exponential tails. Thus, by Eqs. (B2), (B3) and (B4), $\mathrm{VIX}_{t,\Delta,h}(\boldsymbol{\theta}_0, \boldsymbol{\phi}_0, \boldsymbol{\lambda}_0)$ is also geometrically β-mixing with exponential tails. Therefore, $\mathrm{VIX}_{t,\Delta,h}(\boldsymbol{\theta}_0, \boldsymbol{\phi}_0, \boldsymbol{\lambda}_0)$ has enough finite moments to satisfy sufficient conditions for the law of large numbers and the central limit theorem to apply. Next, note that $\mathrm{VIX}_{t,\Delta,h}(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda})$ is continuously differentiable in the interior of $\boldsymbol{\Phi}_0 \times \boldsymbol{\Theta}_0 \times \boldsymbol{\Lambda}_0$ and, hence, the uniform law of large numbers also applies. We may now proceed with a proof that follows the general lines of Propositions 2 and 3, provided we take into account the contribution of parameter estimation error, arising because the risk-neutral paths of the factors are generated using $\hat{\boldsymbol{\phi}}_T$ and $\hat{\boldsymbol{\theta}}_T$, not the unknown $\boldsymbol{\phi}_0$ and $\boldsymbol{\theta}_0$.

Given the first order conditions, and a mean value expansion around $\boldsymbol{\lambda}_0$,

$$\begin{aligned}
&\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\lambda}}_{\mathcal{T}} - \boldsymbol{\lambda}_0\right) \\
&= -\left(\nabla_\lambda \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\lambda}}_{\mathcal{T}})\right)^{\top} \nabla_\lambda \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \bar{\boldsymbol{\lambda}}_{\mathcal{T}})\right)\right)^{-1} \\
&\quad \times \nabla_\lambda \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\lambda}}_{\mathcal{T}})\right)^{\top} \sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \boldsymbol{\lambda}_0) - \tilde{\boldsymbol{\psi}}_{\mathcal{T}}\right),
\end{aligned}$$

where $\bar{\boldsymbol{\lambda}}_{\mathcal{T}}$ is some convex combination of $\hat{\boldsymbol{\lambda}}_{\mathcal{T}}$ and $\boldsymbol{\lambda}_0$. Let

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

Given Propositions 2 and 3, by the uniform law of large numbers, we have that $\hat{\boldsymbol{\lambda}}_{\mathcal{T}} - \boldsymbol{\lambda}_0 = o_p(1)$ and, also, $\sup_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}_0, \boldsymbol{\theta}\in\boldsymbol{\Theta}_0, \psi\in\Psi_0} \left|\nabla_\lambda \left(\tfrac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\lambda})\right) - \boldsymbol{D}_3(\boldsymbol{\phi}, \boldsymbol{\theta}, \boldsymbol{\lambda})\right| = o_p(1)$. Therefore, we have that $\nabla_\lambda \left(\tfrac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\lambda}}_{\mathcal{T}})\right) - \boldsymbol{D}_3 = o_p(1)$, and, hence,

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

We have,

$$\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \boldsymbol{\lambda}_0) - \boldsymbol{\psi}_0\right)$$

$$
\begin{aligned}
&= \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,\boldsymbol{\lambda}_0\right)-\boldsymbol{\psi}_0\right)+\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T,\hat{\boldsymbol{\theta}}_T,\boldsymbol{\lambda}_0)-\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0)\right)\right)\\
&+\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0)-\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\boldsymbol{\phi}_0,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0\right)\right)\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,\boldsymbol{\lambda}_0\right)-\boldsymbol{\psi}_0\right)+\nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T,\bar{\boldsymbol{\theta}}_T,\boldsymbol{\lambda}_0)\right)\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\theta}}_T-\boldsymbol{\theta}_0\right)\\
&+\nabla_\phi\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\bar{\boldsymbol{\phi}}_T,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0\right)\right)\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\phi}}_T-\boldsymbol{\phi}_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,\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)+o_p(1),
\end{aligned}
$$

where $\pi=\lim_{T,\mathcal{T}\to\infty}\mathcal{T}/T$, $\bar{\boldsymbol{\theta}}_T$ and $\bar{\boldsymbol{\phi}}_T$ are convex combinations of $\hat{\boldsymbol{\theta}}_T,\boldsymbol{\theta}_0$, and $\hat{\boldsymbol{\phi}}_T,\boldsymbol{\phi}_0$, respectively, and:

$$
\boldsymbol{F}_{\theta_0}^{\top}=\underset{T,\mathcal{T}\to\infty}{\text{plim}}\nabla_\theta\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\hat{\boldsymbol{\phi}}_T,\bar{\boldsymbol{\theta}}_T,\boldsymbol{\lambda}_0)\right)\quad\text{and}\quad\boldsymbol{F}_{\phi_0}^{\top}=\underset{T,\mathcal{T}\to\infty}{\text{plim}}\nabla_\phi\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}(\bar{\boldsymbol{\phi}}_T,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0)\right).
$$

Therefore, we have:

$$
\text{Avar}\left(\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\lambda}}_{\mathcal{T}}-\boldsymbol{\lambda}_0\right)\right)=\left(\boldsymbol{D}_3^{\top}\boldsymbol{D}_3\right)^{-1}\boldsymbol{D}_3^{\top}\boldsymbol{J}_{00}\boldsymbol{D}_3\left(\boldsymbol{D}_3^{\top}\boldsymbol{D}_3\right)^{-1},
$$

where

$$
\begin{aligned}
\boldsymbol{J}_{00}&=\text{Avar}\left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\boldsymbol{\phi}_0,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0\right)-\boldsymbol{\psi}_0\right)\right)+\text{Avar}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right)\\
&-2\text{Acov}\left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\boldsymbol{\phi}_0,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0\right)-\boldsymbol{\psi}_0\right),\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right)\\
&+\pi\boldsymbol{F}_{\theta_0}^{\top}\text{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T-\boldsymbol{\theta}_0\right)\right)\boldsymbol{F}_{\theta_0}+\pi\boldsymbol{F}_{\phi_0}^{\top}\text{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T-\boldsymbol{\phi}_0\right)\right)\boldsymbol{F}_{\phi_0}\\
&+2\pi\text{Acov}\left(\boldsymbol{F}_{\phi_0}^{\top}\sqrt{T}\left(\hat{\boldsymbol{\phi}}_T-\boldsymbol{\phi}_0\right),\boldsymbol{F}_{\theta_0}^{\top}\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T-\boldsymbol{\theta}_0\right)\right)\\
&+2\sqrt{\pi}\text{Acov}\left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\boldsymbol{\phi}_0,\boldsymbol{\theta}_0,\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}\text{Acov}\left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{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}\text{Acov}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{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}\text{Acov}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right),\boldsymbol{F}_{\theta_0}^{\top}\sqrt{T}\left(\hat{\boldsymbol{\theta}}_T-\boldsymbol{\theta}_0\right)\right).
\end{aligned}
$$

Let $\hat{\boldsymbol{\psi}}_{\mathcal{T},h}(\hat{\boldsymbol{\phi}}_T,\hat{\boldsymbol{\theta}}_T,\boldsymbol{\lambda}_0)$ be the estimator obtained in the case we computed the model-based VIX using simulated paths for the unobservable factor $Z_{t,\Delta,h}^{\theta_u}$. By an argument similar to that in Proposition 2,

$$
\text{Avar}\left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},\Delta,h}\left(\boldsymbol{\phi}_0,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0\right)-\boldsymbol{\psi}_0\right)\right)=\text{Avar}\left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},h}\left(\boldsymbol{\phi}_0,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0\right)-\boldsymbol{\psi}_0\right)\right).
$$

By the same argument as in Proposition 3,

$$
\begin{aligned}
&\text{Avar}\left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},h}\left(\boldsymbol{\phi}_0,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0\right)-\boldsymbol{\psi}_0\right)\right)+\text{Avar}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right)\\
&-2\text{Acov}\left(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\psi}}_{\mathcal{T},h}\left(\boldsymbol{\phi}_0,\boldsymbol{\theta}_0,\boldsymbol{\lambda}_0\right)-\boldsymbol{\psi}_0\right),\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right)=\left(1+\frac{1}{H}\right)\left(\boldsymbol{J}_3-\boldsymbol{K}_3\right),
\end{aligned}
$$

where

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

and

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

Hence, given Propositions 2 and 3,

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

where

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

B.2. Bootstrap estimates of the standard errors

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 re-sampled 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 re-sampled series, $\boldsymbol{X}_t^*$, and let $\mathrm{E}^*,\text{var}^*$ 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.

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 2, the estimators of the auxiliary model parameters, based on actual and simulated samples, have the same asymptotic variance. Hence, there is no need to re-sample 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 B1: *Under the same assumptions of Proposition 2, if $l/T^{1/2}\to0$ as $T,B,l\to\infty$, then for all $\varepsilon>0$,*

$$\Pr\left(\omega:P^*\left(\left|\left(1+\frac{1}{H}\right)\hat{\boldsymbol{V}}_{1,T,B}-\boldsymbol{V}_1\right|>\varepsilon\right)\right)\to0.$$

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

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

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

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

The Proposition follows, once we show that:

$$\mathrm{E}^* \left(\sqrt{T} \left(\tilde{\boldsymbol{\varphi}}_T^* - \tilde{\boldsymbol{\varphi}}_T \right) \right) = o_p(1), \tag{B5}$$

$$\mathrm{var}^* \left(\sqrt{T} \left(\tilde{\boldsymbol{\varphi}}_T^* - \tilde{\boldsymbol{\varphi}}_T \right) \right) = \mathrm{var} \left(\sqrt{T} \left(\tilde{\boldsymbol{\varphi}}_T - \boldsymbol{\varphi}_0 \right) \right) + O_p(l/\sqrt{T}), \tag{B6}$$

and for $\varepsilon > 0$,

$$\mathrm{E}^* \left(\left(\sqrt{T} \, \| \tilde{\boldsymbol{\varphi}}_T^* - \tilde{\boldsymbol{\varphi}}_T \| \right)^{2+\varepsilon} \right) = O_p(1). \tag{B7}$$

Indeed, under conditions (B5)-(B6), 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,

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

and, given (B6), and recalling that $l/\sqrt{T} \to 0$,

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

Given (B7), the statement follows by Theorem 1 in Goncalves and White (2005).

Let us show (B5), (B6) and (B7). We have,

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

Since each component of $\sqrt{T} \left(\tilde{\boldsymbol{\varphi}}_T^* - \tilde{\boldsymbol{\varphi}}_T \right)$ 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. (25), and $\boldsymbol{Y}_t^*$ be its bootstrap counterpart. By the first order conditions,

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

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 (B5). Next,

$$\mathrm{var}^* \left(\sqrt{T} \left(\tilde{\boldsymbol{\varphi}}_{1,T}^* - \tilde{\boldsymbol{\varphi}}_{1,T} \right) \right)$$

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

where $\tilde{\epsilon}_{1,t} = y_{1,t} - \boldsymbol{Y}_t^\top\tilde{\boldsymbol{\varphi}}_{1,T}$. This proves (B6). Finally, as $\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t\boldsymbol{Y}_t^\top$ is full rank, by the same argument used above, for a generic constant C, and $\varepsilon > 0$,

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

This proves (B7).

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. Thus, we need to take into account the contribution of $\boldsymbol{K}_2$, the covariance between simulated and sample paths, as well as among paths at different simulation replications.

Construct the re-sampled simulated stock price series as:

$$
s_{t,\Delta,h}^{\theta,*} = s_0 + s_1 y_{1,t}^* + s_2 y_{2,t}^* + Z_{t,\Delta,h}^{\theta_u,*}, \tag{B8}
$$

where $Z_{t,\Delta,h}^{\theta_u,*}$ is re-sampled from the simulated unobservable process $Z_{t,\Delta,h}^{\theta_u}$, and use $s_{t,\Delta,h}^{\theta,*}$ to construct $R_{t,\Delta,h}^*(\boldsymbol{\theta})$ and $\mathrm{Vol}_{t,\Delta,h}^*(\boldsymbol{\theta})$. 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 re-sampled observations, and $\overline{R}^*$, $\overline{\mathrm{Vol}}^*$ are the sample means of $R_t^* = \ln(S_t^*/S_{t-12}^*)$ and $\mathrm{Vol}_t^* = \sqrt{6\pi}\cdot\frac{1}{12}\sum_{i=1}^{12}|\ln(S_{t+1-i}^*/S_{t-i}^*)|$, with S_t^* being the re-sampled series of the observable stock price process S_t, and

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

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

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

where $\hat{\boldsymbol{\vartheta}}_{T,\Delta,h,i}^*(\boldsymbol{\theta})$ and $\tilde{\boldsymbol{\vartheta}}_{T,i}^*$ denote the values of $\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}^*(\boldsymbol{\theta})$ and $\tilde{\boldsymbol{\vartheta}}_T^*$ a t 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 $\left(1+\frac{1}{H}\right)\hat{\boldsymbol{V}}_{2,T,B}$ is a consistent estimator of $\boldsymbol{V}_2$, and can then be used to obtain asymptotically valid bootstrap standard errors.

Proposition B2: *Under the same assumptions of Proposition 3, if $l/T^{1/2} \to 0$ as $T, B, l \to \infty$, then, for all $\varepsilon > 0$,*

$$
\Pr\left(\omega : P^*\left(\left|\left(1+\frac{1}{H}\right)\hat{\boldsymbol{V}}_{2,T,B} - \boldsymbol{V}_2\right| > \varepsilon\right)\right) \to 0.
$$

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

$$
0 = \nabla_\theta\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}^*(\hat{\boldsymbol{\theta}}_T^*)\right)^\top\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\left(\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}^*(\hat{\boldsymbol{\theta}}_T^*) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right) - \left(\tilde{\boldsymbol{\vartheta}}_T^* - \tilde{\boldsymbol{\vartheta}}_T\right)\right)
$$

$$= \nabla_\theta\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}^*_T)\right)^\top \left(\tfrac{1}{H}\textstyle\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right) - \left(\tilde{\boldsymbol{\vartheta}}^*_T - \tilde{\boldsymbol{\vartheta}}_T\right)\right)$$
$$+ \nabla_\theta\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}^*_T)\right)^\top \nabla_\theta\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}\left(\bar{\boldsymbol{\theta}}^*_T\right)\right)\left(\hat{\boldsymbol{\theta}}^*_T - \hat{\boldsymbol{\theta}}_T\right),$$

where $\bar{\boldsymbol{\theta}}^*_T$ is a convex combination of $(\hat{\boldsymbol{\theta}}^*_T, \hat{\boldsymbol{\theta}}_T)$. Hence,

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}^*_T - \hat{\boldsymbol{\theta}}_T\right) = -\left(\nabla_\theta\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}^*_T)\right)^\top \nabla_\theta\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}\left(\bar{\boldsymbol{\theta}}^*_T\right)\right)\right)^{-1}$$
$$\times\nabla_\theta\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}^*_T)\right)^\top \sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right) - \left(\tilde{\boldsymbol{\vartheta}}^*_T - \tilde{\boldsymbol{\vartheta}}_T\right)\right).$$

We need to show that:

$$\mathrm{E}^*\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right) = o_p(1), \tag{B9}$$

$$\mathrm{var}^*\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right) = \mathrm{var}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \vartheta\left(\boldsymbol{\theta}_0\right)\right)\right)\right) + o_p(1), \tag{B10}$$

and for all $\varepsilon > 0$,

$$\mathrm{E}^*\left(\left\|\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right\|^{2+\varepsilon}\right) < \infty. \tag{B11}$$

The statement in the Proposition follows by the same argument as that in the proof of Proposition B2. Note that,

$$\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right) = \begin{pmatrix} \sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{1,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{1,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right) \\ \sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right) \\ \sqrt{T}\left(\bar{R}^*_{\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \bar{R}_{\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right) \\ \sqrt{T}\left(\overline{\mathrm{Vol}}^*_{\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \overline{\mathrm{Vol}}_{\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right) \end{pmatrix}.$$

We only consider $\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)$, as the remaining terms can be dealt with in the same manner. Let $\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)$ be the vector containing all the simulated regressors in Eq. (31). This vector depends on $\hat{\boldsymbol{\theta}}_T$ because it includes the simulated volatility, obtained with parameter vector $\boldsymbol{\theta}$ fixed at its estimate, $\hat{\boldsymbol{\theta}}_T$. Likewise, let $\boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)$ be the vector containing the bootstrap values of all the simulated regressors in Eq. (31). This vector depends on $\hat{\boldsymbol{\theta}}_T$ because it includes the bootstrapped value of the simulated volatility in Eq. (31), which we denote with $\mathrm{Vol}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)$. By the first order conditions,

$$\mathrm{E}^*\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H\left(\hat{\boldsymbol{\vartheta}}^*_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right)$$
$$= \frac{1}{H}\sum_{h=1}^H \mathrm{E}^*\left(\left(\tfrac{1}{T}\textstyle\sum_{t=13}^T \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\right)^{-1}\tfrac{1}{\sqrt{T}}\textstyle\sum_{t=13}^T \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\left(\mathrm{Vol}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)$$
$$= \frac{1}{H}\sum_{h=1}^H\left(\tfrac{1}{T}\textstyle\sum_{t=13}^T \boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\right)^{-1}\mathrm{E}^*\left(\tfrac{1}{\sqrt{T}}\textstyle\sum_{t=13}^T \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\left(\mathrm{Vol}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)$$
$$+ \frac{1}{H}\sum_{h=1}^H \mathrm{E}^*\left(\left(\left(\tfrac{1}{T}\textstyle\sum_{t=13}^T \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\right)^{-1} - \left(\tfrac{1}{T}\textstyle\sum_{t=13}^T \boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\right)^{-1}\right)\right.$$
$$\left.\times\left(\tfrac{1}{\sqrt{T}}\textstyle\sum_{t=13}^T \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\left(\mathrm{Vol}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right) = \mathrm{E}^*\left(I^*_{T,h}\right) + \mathrm{E}^*\left(II^*_{T,h}\right),$$

We have,

$$\mathrm{E}^*\left(I^*_{T,h}\right)$$
$$= \frac{1}{H}\sum_{h=1}^H\left(\tfrac{1}{T}\textstyle\sum_{t=13}^T \boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\right)^{-1}\mathrm{E}^*\left(\tfrac{1}{\sqrt{T}}\textstyle\sum_{t=13}^T \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\left(\mathrm{Vol}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)$$

$$= O_p\left(l/\sqrt{T}\right) = o_p(1),$$

and as $II^*_{T,h}$ is of smaller order that $I^*_{T,h}$, $\mathrm{E}^*(II^*_{T,h}) = o_p(1)$. This proves (B9). Next, we have that for $h = 1, \cdots, H$,

$$\begin{aligned}\mathrm{E}^*\left(\left(\frac{1}{T}\sum_{t=13}^{T}\boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\right)^{-1}\right) &= \left(\frac{1}{T}\sum_{t=13}^{T}\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\right)^{-1} + o_p(1) \\ &= \left(\mathrm{E}\left(\boldsymbol{U}_{t,\Delta,h}\left(\boldsymbol{\theta}_0\right)\boldsymbol{U}_{t,\Delta,h}\left(\boldsymbol{\theta}_0\right)^\top\right)\right)^{-1} + o_p(1).\end{aligned}$$

Therefore, we need to show that

$$\begin{aligned} & \mathrm{var}^*\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\left(\mathrm{Vol}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right) \\ = \; & \mathrm{Avar}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\left(\mathrm{Vol}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right) \\ & + o_p(1).\end{aligned}$$

Because the blocks are all independent,

$$\begin{aligned} & \mathrm{var}^*\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\left(\mathrm{Vol}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}^*_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right) \\ & = \frac{1}{T}\frac{1}{H^2}\sum_{h=1}^{H}\sum_{h'=1}^{H}\sum_{j=13-l}^{l}\sum_{t=13+l}^{T-l}\hat{\epsilon}_{t,\Delta,h}\hat{\epsilon}_{t+j,\Delta,h'}\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\boldsymbol{U}_{t+j,\Delta,h'}(\hat{\boldsymbol{\theta}}_T)^\top + o_p(1) \\ & = \mathrm{Avar}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\boldsymbol{U}_{t,\Delta,h}\left(\boldsymbol{\theta}_0\right)\left(\mathrm{Vol}_{t,\Delta,h}(\boldsymbol{\theta}_0) - \boldsymbol{U}_{t,\Delta,h}\left(\boldsymbol{\theta}_0\right)^\top\boldsymbol{\vartheta}_{2,0}\left(\boldsymbol{\theta}_0\right)\right)\right)\right) \\ & + o_p(1),\end{aligned}$$

where $\hat{\epsilon}_{t,\Delta,h} = \mathrm{Vol}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)$. This proves Eq. (B10). Finally, under the parameter restrictions in Assumptions B1(i) and B1(iii),

$$\begin{aligned} & \mathrm{E}^*\left(\left\|\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\boldsymbol{\vartheta}}^*_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right\|^{2+\varepsilon}\right) \\ & = \left\|\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\left(\mathrm{Vol}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T) - \boldsymbol{U}_{t,\Delta,h}(\hat{\boldsymbol{\theta}}_T)^\top\hat{\boldsymbol{\vartheta}}_{2,T,\Delta,h}(\hat{\boldsymbol{\theta}}_T)\right)\right)\right\|^{2+\varepsilon} \\ & = O_p(1).\end{aligned}$$

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 re-sample $y_{1,t}, y_{2,t}, S_t$ and VIX_t from the shorter sample, using blocksize l and number of blocks B, so that $lB = \mathcal{T}$. Also, we need to re-sample the unobservable factor from a sample of length $\mathcal{T}$, at the parameter estimate of $\boldsymbol{\theta}_u$ obtained in the previous step, $\hat{Z}^{\theta_u}_{t,\Delta,h}$ say. Let $\mathrm{VIX}^*_{t,\Delta,h}(\boldsymbol{y}^*_t;\hat{\boldsymbol{\phi}}^*_T,\hat{\boldsymbol{\theta}}^*_T,\boldsymbol{\lambda})$ be the model-based VIX index constructed using $y^*_{1,t}$, $y^*_{2,t}$ and the unobservable factor re-sampled at the bootstrap estimators $\hat{\boldsymbol{\phi}}^*_T$ and $\hat{\boldsymbol{\theta}}^*_T$, $\hat{Z}^{\theta_u,*}_{t,\Delta,h}$ say. Finally, let

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

where $\tilde{\boldsymbol{\psi}}^*_{1,\mathcal{T}}$ are the auxiliary model parameters estimated using $y^*_{1,t}, y^*_{2,t}$, and VIX^*_t, with VIX^*_t being the re-sampled 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:

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

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

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

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 B1 and B2. It shows that $\left(1+\frac{1}{H}\right)\hat{\boldsymbol{V}}_{3,\mathcal{T},B}$ is a consistent estimator of $\boldsymbol{V}_3$, and can then provide asymptotically valid bootstrap standard errors.

Proposition B3: *Under the same assumptions of Proposition 4, if $l/\mathcal{T}^{1/2} \to 0$ as $T, \mathcal{T}, B, l \to \infty$, then, for all $\varepsilon > 0$,*

$$\Pr\left(\omega : P^* \left(\left| \left(1+\frac{1}{H}\right) \hat{\boldsymbol{V}}_{3,\mathcal{T},B} - \boldsymbol{V}_3 \right| > \varepsilon \right) \right) \to 0.$$

Proof: We drop the subscript Δ, since for $\Delta\sqrt{T} \to 0$, $\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\psi}}^*_{\mathcal{T},\Delta,h}(\boldsymbol{\phi},\boldsymbol{\theta},\boldsymbol{\lambda}) - \hat{\boldsymbol{\psi}}^*_{\mathcal{T},h}(\boldsymbol{\phi},\boldsymbol{\theta},\boldsymbol{\lambda}) \right) = o^*_p(1)$. By the first order conditions and a mean value expansion around $\hat{\boldsymbol{\lambda}}_{\mathcal{T}}$,

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

where $\bar{\boldsymbol{\lambda}}^*_{\mathcal{T}}$ is some convex combination of $\hat{\boldsymbol{\lambda}}^*_{\mathcal{T}}$ and $\hat{\boldsymbol{\lambda}}_{\mathcal{T}}$. We have:

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

By the same arguments used to show Propositions B1 and B2, and the condition that $\mathcal{T}/T \to \pi \in (0,1)$,

$$\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\phi}}^*_T - \boldsymbol{\phi}_0 \right) = \sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\phi}}^*_T - \hat{\boldsymbol{\phi}}_T \right) + \sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\phi}}_T - \boldsymbol{\phi}_0 \right) = O^*_p(1),$$

and

$$\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\theta}}^*_T - \boldsymbol{\theta}_0 \right) = \sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\theta}}^*_T - \hat{\boldsymbol{\theta}}_T \right) + \sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0 \right) = O^*_p(1).$$

Therefore, by the uniform law of large numbers,

$$\hat{\boldsymbol{\lambda}}^*_{\mathcal{T}} - \boldsymbol{\lambda}_0 = \left(\hat{\boldsymbol{\lambda}}^*_{\mathcal{T}} - \hat{\boldsymbol{\lambda}}_{\mathcal{T}} \right) + \left(\hat{\boldsymbol{\lambda}}_{\mathcal{T}} - \boldsymbol{\lambda}_0 \right) = o^*_p(1) + o_p(1) = o^*_p(1)$$

and, hence,

$$\nabla_\lambda \left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H} \hat{\boldsymbol{\psi}}^*_{\mathcal{T},h}(\hat{\boldsymbol{\phi}}^*_T, \hat{\boldsymbol{\theta}}^*_T, \hat{\boldsymbol{\lambda}}^*_{\mathcal{T}}) \right) - \boldsymbol{D}_3 = o^*_p(1).$$

By an argument similar to that in the proof of Proposition B2,

$$\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H} \left(\hat{\boldsymbol{\psi}}^*_{\mathcal{T},h}(\hat{\boldsymbol{\phi}}^*_T, \hat{\boldsymbol{\theta}}^*_T, \hat{\boldsymbol{\lambda}}_{\mathcal{T}}) - \hat{\boldsymbol{\psi}}_{\mathcal{T},h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\lambda}}_{\mathcal{T}}) \right) \right)$$

$$
\begin{aligned}
=&\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\boldsymbol{\psi}}^{*}_{\mathcal{T},h}(\hat{\boldsymbol{\phi}}_{T},\hat{\boldsymbol{\theta}}_{T},\hat{\boldsymbol{\lambda}}_{\mathcal{T}})-\hat{\boldsymbol{\psi}}_{\mathcal{T},h}(\hat{\boldsymbol{\phi}}_{T},\hat{\boldsymbol{\theta}}_{T},\hat{\boldsymbol{\lambda}}_{\mathcal{T}})\right)\right)+\frac{1}{H}\sum_{h=1}^{H}\nabla_{\phi}\left(\hat{\boldsymbol{\psi}}^{*}_{\mathcal{T},h}(\bar{\boldsymbol{\phi}}^{*}_{T},\hat{\boldsymbol{\theta}}^{*}_{T},\hat{\boldsymbol{\lambda}}_{\mathcal{T}})\right)^{\top}\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\phi}}^{*}_{T}-\hat{\boldsymbol{\phi}}_{T}\right)\\
&+\frac{1}{H}\sum_{h=1}^{H}\nabla_{\theta}\left(\hat{\boldsymbol{\psi}}^{*}_{\mathcal{T},h}(\hat{\boldsymbol{\phi}}_{T},\bar{\boldsymbol{\theta}}^{*}_{T},\hat{\boldsymbol{\lambda}}_{\mathcal{T}})\right)^{\top}\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\theta}}^{*}_{T}-\hat{\boldsymbol{\theta}}_{T}\right)+o^{*}_{p}(1)\\
=&\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\boldsymbol{\psi}}^{*}_{\mathcal{T},h}\left(\boldsymbol{\phi}_{0},\boldsymbol{\theta}_{0},\boldsymbol{\lambda}_{0}\right)-\hat{\boldsymbol{\psi}}_{\mathcal{T},h}(\boldsymbol{\phi}_{0},\boldsymbol{\theta}_{0},\boldsymbol{\lambda}_{0})\right)\right)+\sqrt{\mathcal{T}}\boldsymbol{F}^{\top}_{\phi_0}\left(\hat{\boldsymbol{\phi}}^{*}_{T}-\hat{\boldsymbol{\phi}}_{T}\right)+\sqrt{\mathcal{T}}\boldsymbol{F}^{\top}_{\theta_0}\left(\hat{\boldsymbol{\theta}}^{*}_{T}-\hat{\boldsymbol{\theta}}_{T}\right)+o^{*}_{p}(1).
\end{aligned}
$$

Therefore, by the same argument used in the proof of Propositions B1 and B2, we can show that

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

and

$$
\begin{aligned}
&\mathrm{var}^{*}\left(\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\lambda}}^{*}_{\mathcal{T}}-\hat{\boldsymbol{\lambda}}_{\mathcal{T}}\right)\right)\\
&=\mathrm{var}^{*}\Bigg(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\left(\hat{\boldsymbol{\psi}}^{*}_{\mathcal{T},h}\left(\boldsymbol{\phi}_{0},\boldsymbol{\theta}_{0},\boldsymbol{\lambda}_{0}\right)-\hat{\boldsymbol{\psi}}_{\mathcal{T},h}\left(\boldsymbol{\phi}_{0},\boldsymbol{\theta}_{0},\boldsymbol{\lambda}_{0}\right)\right)\right)\\
&\qquad+\sqrt{\mathcal{T}}\boldsymbol{F}^{\top}_{\phi_0}\left(\hat{\boldsymbol{\phi}}^{*}_{T}-\hat{\boldsymbol{\phi}}_{T}\right)+\sqrt{\mathcal{T}}\boldsymbol{F}^{\top}_{\theta_0}\left(\hat{\boldsymbol{\theta}}^{*}_{T}-\hat{\boldsymbol{\theta}}_{T}\right)-\left(\tilde{\boldsymbol{\psi}}^{*}_{\mathcal{T}}-\tilde{\boldsymbol{\psi}}_{\mathcal{T}}\right)\Bigg)\\
&=\mathrm{Avar}\Bigg(\sqrt{\mathcal{T}}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^{H}\left(\hat{\boldsymbol{\psi}}_{\mathcal{T},h}\left(\boldsymbol{\phi}_{0},\boldsymbol{\theta}_{0},\boldsymbol{\lambda}_{0}\right)-\boldsymbol{\psi}_{0}\left(\boldsymbol{\phi}_{0},\boldsymbol{\theta}_{0},\boldsymbol{\lambda}_{0}\right)\right)\right)\\
&\qquad+\sqrt{\mathcal{T}}\boldsymbol{F}^{\top}_{\phi_0}\left(\hat{\boldsymbol{\phi}}_{T}-\boldsymbol{\phi}_{0}\right)+\sqrt{\mathcal{T}}\boldsymbol{F}^{\top}_{\theta_0}\left(\hat{\boldsymbol{\theta}}_{T}-\boldsymbol{\theta}_{0}\right)-\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_{0}\right)\Bigg)+o_{p}(1).
\end{aligned}
$$

and by Minkowski's inequality, $\mathrm{E}^{*}\left(\left\|\sqrt{\mathcal{T}}\left(\hat{\boldsymbol{\lambda}}^{*}_{\mathcal{T}}-\hat{\boldsymbol{\lambda}}_{\mathcal{T}}\right)\right\|^{2+\varepsilon}\right)=O^{*}_{p}(1)$, for some $\varepsilon>0$.