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

Valentina Corradi
University of Warwick

Walter Distaso
Imperial College Business School

Antonio Mele
London School of Economics

First draft: July 22, 2005. This version: January 5, 2009.

Abstract

This paper introduces a no-arbitrage framework to assess how macroeconomic factors help explain the risk-premium agents require to bear the risk of fluctuations in stock market volatility. We develop a model in which stock volatility and volatility risk-premia are stochastic and derive no-arbitrage conditions linking volatility to macroeconomic factors. We estimate the model using data related to variance swaps, which are contracts with payoffs indexed to nonparametric measures of realized volatility. We find that volatility risk-premia are strongly countercyclical, even more so than standard measures of return volatility.

Keywords: volatility risk-premium; macroeconomic factors; no arbitrage restrictions; indirect inference.

*We thank Yacine Aït-Sahalia, Marcelo Fernandes, Christian Julliard, Mark Salmon, George Tauchen, Viktor Todorov, and seminar participants at 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. 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, return volatility is exogenously driven by some unobservable factor correlated with the asset returns. Yet such an unobservable factor does not bear a direct economic interpretation. Moreover, the model implies, by assumption, that volatility can not 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, with volatility being higher during recessions than during expansions; see, e.g., Schwert (1989a and 1989b) and Brandt and Kang (2004).

In this paper, we develop a no-arbitrage model in which stock market volatility is explicitly related to a number of macroeconomic and unobservable factors. The distinctive feature of the model is that return 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 volatility and volatility-related risk-premia change in response to business cycle conditions. Our focus on the volatility risk-premium is related to the seminal work of Britten-Jones and Neuberger (2000), which has more recently stimulated an increasing interest in the study of the dynamics and determinants of the volatility risk-premium (see, for example, Bakshi and Madan (2006) and Carr and Wu (2007)). 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. In previous important work, Bollerslev, Gibson and Zhou (2004) and Bollerslev and Zhou (2005) unveil, empirically, a strong relation between this volatility risk-premium and a number of macroeconomic factors. In this paper, we make a step further and make the volatility risk-premium be endogenously determined within our no-arbitrage model. The resulting relation between the volatility risk-premium and the macroeconomic factors is richer than in previous empirical investigations, as we are explicitly accounting for the necessary no-arbitrage relations that link asset prices and, hence, return volatility, to macroeconomic factors. More generally, our paper is the first to formulate and estimate a model that relates the dynamics of volatility and volatility adjustments to macroeconomic factors, within a fully-specified no-arbitrage setup. 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 years, there has been an important surge of interest in general equilibrium (GE, henceforth) models linking aggregate stock market volatility to variations in the key factors tracking the state of the economy (see, for example, Campbell and Cochrane (1999), Bansal and Yaron (2004), Mele (2007), and Tauchen (2005)). These GE models are important as they highlight the main economic mechanisms through which markets, preferences and technology affect the equilibrium asset prices and, hence, return volatility. At the same time, we do not observe the emergence of a well accepted paradigm. Rather, a variety of GE models aim to explain the stylized features of aggregate stock market fluctuations (see, for example, Campbell (2003) and Mehra and Prescott (2003) for two views on these issues). In this paper, we do not develop a fully articulated GE model. In our framework, cross-equations restrictions arise through the weaker requirement of absence of arbitrage opportunities. This makes our approach considerably more flexible than it would be under a fully articulated GE discipline. In this respect, our approach is closer in spirit to the "no-arbitrage" vector autoregressions introduced in the term-structure literature by Ang and Piazzesi (2003) and Ang, Piazzesi and Wei (2005). Similarly as in these papers, we specify an analytically convenient pricing kernel affected by some macroeconomic factors, although we do not directly related these to markets, preferences and technology.

Our model works quite simply. We start with exogenously specifying the joint dynamics of a number of macroeconomic and latent factors. Then, we assume that dividends and the risk-premia 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, which is set in continuous-time, does not allow for jumps or related market micro-structure effects. Rather, our continuous-time setup leads to an analytically convenient framework that we use to model low frequency

[1]Our model differs from previous formulations such as that in Bekaert and Grenadier (2001), Ang and Liu (2004) or Mamaysky (2002). For 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.

movements in asset volatility and the volatility risk-premium, through the use of macroeconomic and unobservable factors. Carr and Wu (2007), Todorov (2007) and Tauchen and Todorov (2008) do allow for the presence of jumps, although they are not concerned with the cross-equation restrictions relating the volatility risk-premium to state variables driving aggregate low frequency stock market fluctuations which, instead, constitute the central topic of our paper.

The estimation of our models entails a few challenges. In our model, volatility is endogenous, which makes parameters' identification a quite delicate issue. The main difficulty we face is that return volatility arises out of no-arbitrage restrictions. Therefore, all the factors affecting the aggregate stock market also affect stock market volatility. In the 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 particular, our model predicts that return volatility 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. Thus, the advantage of our model (to generate, endogenously, stock market volatility) also brings an identification issue. We address this identification issue by exploiting derivative price 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 by the CBOE since 2003 as the new VIX index. (The CBOE has re-calculated the new VIX index back to 1990.) These data allow us to identify the model.

We implement a two-stage estimation procedure. In the first step, we use data on a broad stock market index and two macroeconomic factors, inflation and industrial production, and estimate all the parameters, taking the parameters related to risk-premia adjustments as given. We implement this step by matching moments related to ex-post stock market returns, realized return volatility and the two macroeconomic factors. In the second step, we use data on the new VIX index, and the two macroeconomic factors, to estimate the risk-premia parameters. In this second step, we implement consistent estimators of the VIX index. Note, the two-stage estimation procedure entails parameter estimation error. To implement an efficient estimator, then, we rely on the block bootstrap of the entire procedure.

The remainder of the paper is organized as follows. In Section 2 we develop a no-arbitrage model for the stock price, return volatility and the volatility risk-premium. Section 3 illustrates the estimation strategy. Section 4 presents our empirical results. Section 5 concludes, and the appendix provides technical 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. We assume 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,

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

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:

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

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

2.2 Risk-premia 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 the development of macroeconomic conditions, through Eq. (2). We assume that asset prices respond passively to movements in the factors affecting macroeconomic conditions.[2] Formally, 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

[2] For analytical convenience, we are ruling out that asset prices can feed back the real economy, although we acknowledge that financial frictions can make financial markets and the macroeconomy intimately related, as in the financial accelerator hypothesis reviewed by Bernanke, Gertler and Gilchrist (1999).

continuously differentiable in $\boldsymbol{y}$. (Given the assumptions and conditions we give below, this differentiability condition holds in our model.) 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 $s_y(y) = [\frac{\partial}{\partial y_1}s(\boldsymbol{y}), \cdots, \frac{\partial}{\partial y_n}s(\boldsymbol{y})]^\top$ and m is a function we shall determine below by no-arbitrage conditions. By Eq. (3), the instantaneous return variance 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 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 Q, the risk-neutral probability measure with respect to P on $\mathbb{F}(T)$ is,

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

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,

$$\frac{\mathrm{d}s(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))} = (r - \delta(\boldsymbol{y}(t)))\,\mathrm{d}t + \frac{s_y(\boldsymbol{y}(t))^\top \boldsymbol{\Sigma V}(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))}\mathrm{d}\hat{\boldsymbol{W}}(t), \tag{5}$$

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-premia and stochastic volatility separately. Given a dividend process, volatility is uniquely determined, once we specify the risk-premia. Consider,

[3] This assumption can be replaced with a weaker condition that the short-term rate is an affine function of the underlying state vector. This assumption would destroy the property that the asset price is affine in the state vector $\boldsymbol{y}$, as established in Proposition 1 below, which would considerably hinder statistical inference.

then, the following "essentially affine" specification for the dynamics of the factors in Eq. (1). Let $\boldsymbol{V}^{-}(\boldsymbol{y})$ be a $d \times d$ diagonal matrix with elements

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

and set,

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

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 in the specification 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-premia $\boldsymbol{\Lambda}$ are tied up to the volatility of the fundamentals, $\boldsymbol{V}(\boldsymbol{y})$. While it is reasonable to assume that risk-premia are related to the *volatility* of fundamentals, the specification in Eq. (6) is more general, as it allows risk-premia 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 developed in the appendix,[4] and assuming no-bubbles, Eq. (5) implies that the stock price is,

$$s(\boldsymbol{y}) = \mathbb{E}\left[\int_0^\infty e^{-rt}\delta(\boldsymbol{y}(t))\,\mathrm{d}t \,\middle|\, \boldsymbol{y}(0) = \boldsymbol{y}\right], \tag{7}$$

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

$$\delta(\boldsymbol{y}) = \delta_0 + \boldsymbol{\delta}^\top \boldsymbol{y}, \tag{8}$$

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

Proposition 1. *Let the risk-premia and the instantaneous dividend rate be as in Eqs. (6) and (8). Then, under a technical regularity condition in the Appendix (condition (A2)), we have that: (i) eq. (7) holds; and (ii) the rational stock price function $s(\boldsymbol{y})$ is linear in the state vector $\boldsymbol{y}$, viz*

$$s(\boldsymbol{y}) = \frac{\delta_0 + \boldsymbol{\delta}^\top (\boldsymbol{D} + r\boldsymbol{I}_{n\times n})^{-1}\boldsymbol{c}}{r} + \boldsymbol{\delta}^\top (\boldsymbol{D} + r\boldsymbol{I}_{n\times n})^{-1}\boldsymbol{y}, \tag{9}$$

[4]These conditions relate to the volatility term $s_y(\boldsymbol{y})^\top \boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})$ in Eq. (3). This term must satisfy integrability conditions ensuring that the Itô's integral in the representation of the discounted stock price is a martingale.

where

$$c = \kappa\mu - \Sigma \begin{pmatrix} \alpha_1\lambda_{1(1)} & \cdots & \alpha_d\lambda_{1(d)} \end{pmatrix}^\top \tag{10}$$

$$D = \kappa + \Sigma \left[\begin{pmatrix} \lambda_{1(1)}\beta_1^\top & \cdots & \lambda_{1(d)}\beta_d^\top \end{pmatrix}^\top + I^- \lambda_2 \right], \tag{11}$$

I^- is a $d \times d$ diagonal matrix with elements $I^-_{(ii)} = 1$ if $\Pr\{V(y(t))_{(ii)} > 0 \text{ all } t\} = 1$ and 0 otherwise; and, finally $\{\lambda_{1(j)}\}_{j=1}^d$ are the components of λ_1.

Proposition 1 allows us to single out the no-arbitrage restrictions between stochastic volatility and risk-premia. In particular, by Eq. (4), and the expression for the stock price in Eq. (9), we have:

$$\sigma(y(t)) \equiv \sigma(t) = \frac{\sqrt{\left\| \delta^\top (D + rI_{n\times n})^{-1} \Sigma V(y(t)) \right\|^2}}{\frac{\delta_0 + \delta^\top (D + rI_{n\times n})^{-1} c}{r} + \delta^\top (D + rI_{n\times n})^{-1} y(t)}. \tag{12}$$

This formula makes clear 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-premia are modeled independently of each other. For example, the celebrated Heston's (1993) model assumes that the stock price is solution to,

$$\begin{cases} \dfrac{ds(t)}{s(t)} = m(t)\,dt + v(t)\,dW_1(t) \\ dv^2(t) = \kappa\left(\mu - v^2(t)\right)dt + \sigma v(t)\left(\rho dW_1(t) + \sqrt{1-\rho^2}dW_2(t)\right) \end{cases} \tag{13}$$

for some adapted process $m(t)$ and some constants $\kappa, \mu, \sigma, \rho$. 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)). 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 (2005)). 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-premia and the business cycle.

Our model works differently because it places restrictions directly on the asset price process, through our assumptions about the fundamentals of the economy, i.e. the dividend process in Eq. (8) and the risk-premia in Eq. (6). In our model, it is the asset price process that determines, endogenously, the volatility dynamics. For this reason, the model predicts that return volatility embeds information about risk-corrections that agents require to invest in the stock market, as Eq. (12) makes clear. We shall make use of this observation in the empirical part of the paper. We now turn to describe which measure of return volatility measure we shall use to proceed with such a critical step of the paper.

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 index return. The new index reflects recent advances into 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. (9)), and all available information $\mathbb{F}(t)$ at time t, define the integrated return variance on a given interval $[t, T]$ as,

$$IV_{t,T} = \int_t^T \mathbb{E}\left[\left(\frac{\mathrm{d}}{\mathrm{d}\tau} var\left[\log s(\tau) \middle| \mathbb{F}(u)\right]\bigg|_{\tau=u}\right)\middle| \mathbb{F}(t)\right] \mathrm{d}u. \tag{14}$$

The new VIX index relies on the work of Bakshi and Madan (2000), Britten-Jones and Neuberger (2000), and Carr and Madan (2001), who showed that the risk-neutral probability expectation of the future integrated variance is a functional of put and call options written on the asset:

$$\mathbb{E}\left[IV_{t,T} \middle| \mathbb{F}(t)\right] = 2e^{-r(T-t)}\left[\int_0^{F(t)} \frac{P(t,T,K)}{K^2}\mathrm{d}K + \int_{F(t)}^{\infty} \frac{C(t,T,K)}{K^2}\mathrm{d}K\right], \tag{15}$$

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, (14), and some strike price, the variance swap rate. In the absence of arbitrage opportunities, then, the variance swap rate is given by Eq. (15).

In contrast, our model predicts that the risk-neutral expectation of the integrated variance is:

$$\mathbb{E}\left[IV_{t,T} \middle| \boldsymbol{y}(t) = \boldsymbol{y}\right] = \int_t^T \mathbb{E}\left[\sigma^2(u) \middle| \boldsymbol{y}(t) = \boldsymbol{y}\right] \mathrm{d}u, \tag{16}$$

where $\sigma^2(t)$ is given in Eq. (12). It is a fundamental objective of this paper to estimate our model so that it predicts a theoretical pattern of the VIX index that matches its empirical counterpart, computed by the CBOE through Eq. (15).[5]

Note that as a by product, we will be able to trace how the volatility risk-premium, defined as,

$$\mathrm{VRP}(\boldsymbol{y}(t)) \equiv \sqrt{\frac{1}{T-t}}\left(\sqrt{\mathbb{E}\left[IV_{t,T} \middle| \boldsymbol{y}(t) = \boldsymbol{y}\right]} - \sqrt{E\left[IV_{t,T} \middle| \boldsymbol{y}(t) = \boldsymbol{y}\right]}\right), \tag{17}$$

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

[5]The VIX index actually relies on a discretized version of Eq. (15), due to the obvious reason that only a finite number of call and put options are available for trading.

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 the stock market development. We define these factors as follows:

$$\log\left(\mathrm{M}_j\left(t\right)/\,\mathrm{M}_j\left(t-12\right)\right)=\log y_j\left(t\right),\quad j=1,2,$$

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. (Data for such macroeconomic aggregates are typically available at a monthly frequency.) Hence, in terms of Eq. (2), the functions $\varphi_j(\boldsymbol{y}) \equiv \log y_j$.

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,

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

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 $\Pr\{\boldsymbol{V}(\boldsymbol{y}(t))_{(ii)}>0 \text{ all } t\}=1$, 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. (6), 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-premia process defined as,

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

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-premia parameters $\lambda_{2(j)}$.

Finally, the instantaneous dividend process $\delta(t)$ in Eq. (8) 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

$$s_0 = \frac{1}{r}\left[\delta_0 + \sum_{j=1}^{3} s_j \left(\kappa_j \mu_j + \bar{\kappa}_j \bar{\mu}_j - \alpha_j \lambda_{1(j)}\right)\right] \tag{22}$$

$$s_j = \frac{\delta_j \left(r + \kappa_i + \lambda_{1(i)}\beta_i + \lambda_{2(i)}\right) - \delta_i \bar{\kappa}_i}{\prod_{h=1}^{2}\left(r + \kappa_h + \lambda_{1(h)}\beta_h + \lambda_{2(h)}\right) - \bar{\kappa}_1 \bar{\kappa}_2}, \quad \text{for } j, i \in \{1,2\} \text{ and } i \neq j \tag{23}$$

$$s_3 = \frac{\delta_3}{r + \kappa_3 + \lambda_{1(3)}\beta_3 + \lambda_{2(3)}} \tag{24}$$

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 can not 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, $\phi^\intercal = \left(\kappa_j, \mu_j, \alpha_j, \beta_j, \bar{\kappa}_j, j = 1, 2\right)$.

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, $\theta^\intercal = (\kappa_3, \mu_3, \alpha_3, \beta_3, s_0, s_j, j = 1, 2, 3)$, while imposing the identifiability condition that $\mu_3 = 1$, as explained below.

In the third step, we estimate the risk premia parameters $\lambda^\intercal = \left(\lambda_{1(1)}, \lambda_{2(1)}, \lambda_{1(2)}, \lambda_{2(2)}, \lambda_{1(3)}, \lambda_{2(3)}\right)$, relying on a functional 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 is then an hybrid of Indirect Inference (Gouriéroux, Monfort and Renault, 1993) and the Simulated Generalized Method of Moments (Duffie and Singleton, 1993), and does not lead to asymptotic efficiency. Alternatively, we might have relied on *efficient* methods, such as the closed-form approximation approach developed by Aït-Sahalia (2008), the efficient simulated nonparametric estimator in Altissimo and Mele (2008), the method of simulated characteristic functions of Carrasco, Chernov, Florens and Ghysels (2007), the simulated nonparametric maximum likelihood approach set forth by Fermanian and Salanié (2004), or the efficient method of moments of Gallant and Tauchen (1996). These methods are suitable to address estimation within a multifactor framework like ours. However, hinging upon these approaches would make the issue of parameter estimation error be considerably more complex, and beyond the scope of this paper.

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 $y_{1,t,\Delta,h}^{\phi}$ and $y_{2,t,\Delta,h}^{\phi}$, where $y_{j,t,\Delta,h}^{\phi}$ is the value at time t taken by the j-th factor, at the h-th simulation performed with the parameter vector ϕ. Then, we estimate the following autoregressive models on both historical and simulated data, for $i = 1, 2$,

$$y_{i,t} = \varphi_{i,0} + \sum_{j\in\{12,24\}} \varphi_{i,1,j} y_{1,t-j} + \sum_{j\in\{12,24\}} \varphi_{i,2,j} y_{2,t-j} + \epsilon_{y,i,t}, \tag{25}$$

and

$$y_{i,t,\Delta,h}^{\phi} = \varphi_{i,0,h} + \sum_{j\in\{12,24\}} \varphi_{i,1,j,h} \cdot y_{1,t-j,\Delta,h}^{\phi} + \sum_{j\in\{12,24\}} \varphi_{i,2,j,h} \cdot y_{2,t-j,\Delta,h}^{\phi} + \epsilon_{y_h,i,t}. \tag{26}$$

Next, let $\tilde{\varphi}_T = \left(\tilde{\varphi}_{1,T}, \tilde{\varphi}_{2,T}, \bar{y}_1, \bar{y}_2, \hat{\sigma}_1, \hat{\sigma}_2\right)^\intercal$ where, for $i = 1, 2$, $\tilde{\varphi}_{1,T}$ and $\tilde{\varphi}_{2,T}$ denote the ordinary least squares estimators of the parameters in Eq. (25) and, for $i = 1, 2$, $\bar{y}_i$ and $\hat{\sigma}_i$ are the sample average and standard deviation of the macroeconomic factors. Likewise, define $\hat{\varphi}_{T,h}^{\Delta}(\phi)$ to be the simulated counterpart to $\tilde{\varphi}_T$ at simulation h, including the ordinary least squares estimator of the parameters in Eq. (26), and the sample averages and standard deviations of the macroeconomic factors.

The estimator of ϕ, the parameters of the process underlying the macroeconomic factors, is:

$$\hat{\phi}_T = \arg\min_{\phi\in\Phi_0} \left\| \frac{1}{H} \sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi) - \tilde{\varphi}_T \right\|^2, \tag{27}$$

where Φ_0 is a compact set of Φ, a parameter set defined in Appendix B.

We have:

Proposition 2: *As $T \to \infty$ and $\Delta\sqrt{T} \to 0$,*

$$\sqrt{T}\left(\hat{\phi}_T - \phi_0\right) \xrightarrow{d} \mathrm{N}\left(0, \mathbf{V}_1\right),$$

where

$$\begin{aligned}
\mathbf{V}_1 &= \left(1 + \frac{1}{H}\right)\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1}\mathbf{D}_1^{\intercal}\mathbf{J}_1\mathbf{D}_1\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1} \\
\mathbf{D}_1 &= \mathrm{p}\lim \nabla_\phi\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}^{\Delta}\left(\phi_0\right)\right) \\
\mathbf{J}_1 &= \mathrm{Avar}\left(\sqrt{T}\left(\tilde{\varphi}_T - \varphi_0\right)\right) = \mathrm{Avar}\left(\sqrt{T}\left(\hat{\varphi}_{T,h}^{\Delta}\left(\phi_0\right) - \varphi_0\right)\right), \textit{ for all } h.
\end{aligned}$$

and ϕ_0 is the minimizer of the moment conditions in Eq. (27) for $T \to \infty$ and $\Delta\sqrt{T} \to 0$.

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. Indeed, by Eq. (21), the parameters s_j are functions of the structural parameters, as established in Eqs. (22)-(24). In particular, we are not able to identify the parameters related to the dividend process and the risk premia parameters: there are many combinations of δ and λ giving rise to the same equilibrium 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 λ 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 simulated series $Z_{t,h}\left(\theta_u\right)$, where $Z_{t,h}$ is the value of the factor at t, at the h-th simulation, when the parameter vector is $\theta_u = (\kappa_3, \alpha_3, \beta_3, s_3)$. Let, then, $s_{t,\Delta,h}^{\theta}$ be the simulation of the stock price process at time t in the h-th simulation, when the parameters are fixed at θ:

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

where we fix the intercept, $s_0 = \bar{s} - s_1\bar{y}_1 - s_2\bar{y}_2 - \bar{Z}_3^{\Delta}(\theta_u)$, and where $\bar{s}$, $\bar{y}_1$, $\bar{y}_2$ and $\bar{Z}_3^{\Delta}(\theta_u)$ are the sample averages of s_t, $y_{1,t}$, $y_{2,t}$ and $Z_{3,\Delta,t}^{\Delta}(\theta_u)$. Note, we simulate the stock price using the *observed* samples for $y_{1,t}$ and $y_{2,t}$, a characteristic that results in improved efficiency, as we shall discuss below.

Following Mele (2007) and Fornari and Mele (2008), we measure the volatility of the monthly continuously compounded price changes, as:

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

Next, define yearly returns as, $R_t = \log(s_t/s_{t-12})$, 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:

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

and

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

Let $\tilde{\vartheta}_T = \left(\tilde{\vartheta}_{1,T}, \tilde{\vartheta}_{2,T}, \bar{R}, \overline{\mathrm{Vol}}\right)^{\intercal}$, where $\bar{R}$ are $\overline{\mathrm{Vol}}$ the sample averages of return and volatility, $\tilde{\vartheta}_{1,T}$ is the ordinary least squares estimate of the parameters in Eq. (30) and $\tilde{\vartheta}_{2,T}$ is the ordinary least squares estimate of the parameters in Eq. (31). Let $\hat{\vartheta}_{T,h}^{\Delta}(\theta)$ be the simulated counterpart to $\tilde{\vartheta}_T$ at simulation h.

The estimator of θ, the vector including the reduced-form parameters s_j and the parameters related to process of the unobservable factor, is:

$$\hat{\theta}_T = \arg\min_{\theta \in \Theta_0} \left\| \frac{1}{H} \sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{\Delta}(\theta) - \tilde{\vartheta}_T \right\|^2, \tag{32}$$

where Θ_0 is a compact set of Θ, a parameter set defined in Appendix B.

We have:

Proposition 3: *As* $T \to \infty$ *and* $\Delta\sqrt{T} \to 0$,

$$\sqrt{T}\left(\hat{\theta}_T - \theta_0\right) \xrightarrow{d} N\left(0, \mathbf{V}_2\right),$$

where

$$\begin{aligned}
\mathbf{V}_2 &= \left(1+\frac{1}{H}\right)\left(\mathbf{D}_2^{\intercal}\mathbf{D}_2\right)^{-1}\mathbf{D}_2^{\intercal}\left(\mathbf{J}_2-\mathbf{K}_2\right)\mathbf{D}_2\left(\mathbf{D}_2^{\intercal}\mathbf{D}_2\right)^{-1} \\
\mathbf{D}_2 &= \mathrm{p}\lim \nabla_\theta \left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}^{\Delta}\left(\theta_0\right)\right) \\
\mathbf{J}_2 &= \mathrm{Avar}\left(\sqrt{T}\left(\tilde{\vartheta}_T-\vartheta_0\right)\right) = \mathrm{Avar}\left(\sqrt{T}\left(\hat{\vartheta}_{T,h}^{\Delta}\left(\theta_0\right)-\vartheta_0\right)\right), \text{ for all } h \\
\mathbf{K}_2 &= \mathrm{Acov}\left(\sqrt{T}\left(\tilde{\vartheta}_T-\vartheta_0\right), \sqrt{T}\left(\hat{\vartheta}_{T,h}^{\Delta}\left(\theta_0\right)-\vartheta_0\right)^{\intercal}\right) \text{ for all } h \\
&= \mathrm{Acov}\left(\sqrt{T}\left(\hat{\vartheta}_{T,h'}^{\Delta}\left(\theta_0\right)-\vartheta_0\right), \sqrt{T}\left(\hat{\vartheta}_{T,h}^{\Delta}\left(\theta_0\right)-\vartheta_0\right)^{\intercal}\right), \text{ for all } h \neq h'.
\end{aligned}$$

and ϕ_0 is the minimizer of the moment conditions in Eq. (32) for $T \to \infty$ and $\Delta\sqrt{T} \to 0$.

Note that the structure of the asymptotic covariance matrix is different from that in Proposition 2. The difference is the presence of the matrix $\mathbf{K}_2$, 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

We are left with estimating the risk-premia parameters, which we do by matching impulse-response functions as well as other sample moments related to the model-free VIX index.

Consider the instantaneous stock volatility predicted by the model, as defined in Eq. (12), $\sigma\left(\boldsymbol{y}\left(t\right)\right)$. The VIX index predicted by the model is,

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

where $\mathbb{E}$ is the expectation under the risk-neutral probability. Although we do not know $\mathrm{VIX}\left(\boldsymbol{y}\right)$ in closed-form, we can make a functional expansion of $\mathbb{E}[\sigma^2\left(\boldsymbol{y}\left(u\right)\right) \mid \boldsymbol{y}\left(t\right)=\boldsymbol{y}]$, as follows,

$$\mathbb{E}[\sigma^2\left(\boldsymbol{y}\left(u\right)\right) \mid \boldsymbol{y}\left(t\right)=\boldsymbol{y}] = \lim_{N\to\infty}\sum_{n=0}^{N}\frac{\left(u-t\right)^n}{n!}\mathcal{A}^n\sigma^2\left(\boldsymbol{y}\right),$$

where $\mathcal{A}$ is the infinitesimal generator under the risk neutral-probability. Hereafter, we set $n=1$, so that

$$\mathrm{VIX}\left(\boldsymbol{y}\left(t\right)\right) = \sqrt{\sigma^2\left(\boldsymbol{y}\left(t\right)\right)+\frac{1}{2}\mathcal{A}\sigma^2\left(\boldsymbol{y}\left(t\right)\right)\left(T-t\right)} \tag{34}$$

where

$$\mathcal{A}\sigma^2(\boldsymbol{y}) = \nabla_{\boldsymbol{y}}\sigma^2(\boldsymbol{y})^\intercal(\boldsymbol{c} - \boldsymbol{D}\boldsymbol{y}) + \frac{1}{2}\left(\sum_{j=1}^{3}(\alpha_j + \beta_j y_j)\nabla_{y_j y_j}\sigma^2(\boldsymbol{y})\right), \tag{35}$$

where the expressions for the Jacobian $\nabla_{\boldsymbol{y}}\sigma^2(\boldsymbol{y})$, the terms $\nabla_{y_j y_j}\sigma^2(\boldsymbol{y})$, and $\boldsymbol{c}$ and $\boldsymbol{D}$ are given in Appendix B.

In the actual computation of Eq. (35), 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 stages: $\hat{\theta}_T$ and $\hat{\phi}_T$. Moreover, we make use of actual samples for the observable factors $y_{1,t}, y_{2,t}$ and simulated samples for the latent factor, where the latter is simulated using the parameters estimated in the second step. Note, given θ and ϕ, we can now identify $\lambda = \left(\lambda_{1(1)}, \lambda_{1(2)}, \lambda_{1(3)}, \lambda_{2(1)}, \lambda_{2(2)}, \lambda_{2(3)}\right)^\intercal$ from $\boldsymbol{c}$ and $\boldsymbol{D}$.

Let VIX_t be the sample time-series for the VIX index, and let $\text{VIX}^{\Delta}_{t,h}(\boldsymbol{y}_t; \hat{\theta}_T, \hat{\phi}_T, \lambda)$ be the model-based VIX index. As the CBOE VIX index is available only since 1990, in this stage we use a sample of length $\mathcal{T}$, with $\mathcal{T} < T$. In the sequel, we rely on the following auxiliary model

$$\text{VIX}_t = a^{\text{VIX}} + \varphi\text{VIX}_{t-1} + \sum_{i\in\{36,48\}} b^{\text{VIX}}_{1,i} y_{1,t-i} + \sum_{i\in\{36,48\}} b^{\text{VIX}}_{2,i} y_{2,t-i} + \epsilon^{\text{VIX}}_t. \tag{36}$$

Define, $\tilde{\psi}_{\mathcal{T}} = \left(\tilde{\psi}_{1,\mathcal{T}}, \overline{\text{VIX}}, \hat{\sigma}_{\text{VIX}}\right)^\intercal$, where $\tilde{\psi}_{1,\mathcal{T}}$ is the ordinary least squares estimator of the parameters in Eq. (36), and $\overline{\text{VIX}}$ and $\hat{\sigma}_{\text{VIX}}$ are the sample average and standard deviation of the VIX index. Likewise, define $\hat{\psi}^{\Delta}_{\mathcal{T},h}(\hat{\theta}_T, \hat{\phi}_T, \lambda)$, the simulated counterpart to $\tilde{\psi}_{\mathcal{T}}$ at simulation h, obtained through simulations of the model-implied index $\text{VIX}^{\Delta}_{t,h}(\boldsymbol{y}_t; \hat{\theta}_T, \hat{\phi}_T, \lambda)$, with $y_{1,t}$ and $y_{2,t}$ fixed at their sample values.

The estimator of λ, the parameters underlying the risk-premium process, is:

$$\hat{\lambda}_{\mathcal{T}} = \arg\min_{\lambda\in\Lambda_0}\left\|\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}(\hat{\phi}_T, \hat{\theta}_T, \lambda) - \tilde{\boldsymbol{\psi}}_{\mathcal{T}}\right\|^2, \tag{37}$$

for some compact set Λ_0.

We have:

Proposition 4: *If for some $\pi \in (0, 1)$, $T, \mathcal{T} \to \infty$, $\Delta\sqrt{T} \to 0$, $\Delta T \to \infty$, $\mathcal{T}/T \to \pi$, then:*

$$\sqrt{\mathcal{T}}\left(\hat{\lambda}_{\mathcal{T}} - \lambda_0\right) \xrightarrow{d} N(0, \mathbf{V}_3),$$

where

$$\mathbf{V}_3 = \left(\mathbf{D}_3^\intercal \mathbf{D}_3\right)^{-1} \mathbf{D}_3^\intercal \left(\left(1+\frac{1}{H}\right)(\mathbf{J}_3 - \mathbf{K}_3) + \mathbf{P}_3\right) \mathbf{D}_3 \left(\mathbf{D}_3^\intercal \mathbf{D}_3\right)^{-1},$$

$$\mathbf{D}_3 = \underset{\mathcal{T}\to\infty}{\mathrm{p}\lim} \nabla_\lambda \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0, \theta_0, \lambda_0\right)\right),$$

$$\mathbf{J}_3 = \mathrm{Avar}\left(\sqrt{\mathcal{T}}\left(\tilde{\psi}_{\mathcal{T}} - \psi_0\right)\right) = \mathrm{Avar}\left(\sqrt{\mathcal{T}}\left(\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right)\right), \textit{ for all } h$$

$$\mathbf{K}_3 = \mathrm{Acov}\left(\sqrt{\mathcal{T}}\left(\left(\tilde{\psi}_{\mathcal{T}} - \psi_0\right)\right), \sqrt{\mathcal{T}}\left(\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right)^\intercal\right) \textit{ for all } h$$

$$= \mathrm{Acov}\left(\sqrt{\mathcal{T}}\left(\hat{\psi}^{\Delta}_{\mathcal{T},h'}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right), \sqrt{\mathcal{T}}\left(\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right)^\intercal\right), \ \forall\, h \neq h'$$

and

$$\begin{aligned}
\mathbf{P}_3 = {} & \pi F_{\theta_0}^\intercal \mathrm{Avar}\left(\sqrt{T}\left(\hat{\theta}_T - \theta_0\right)\right) F_{\theta_0} + \pi F_{\phi_0}^\intercal \mathrm{Avar}\left(\sqrt{T}\left(\hat{\phi}_T - \phi_0\right)\right) F_{\phi_0} \\
& + 2\pi \mathrm{Acov}\left(F_{\phi_0}^\intercal \sqrt{T}\left(\hat{\phi}_T - \phi_0\right), F_{\theta_0}^\intercal \sqrt{T}\left(\hat{\theta}_T - \theta_0\right)\right) \\
& + 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H} \hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right), F_{\phi_0}^\intercal \sqrt{T}\left(\hat{\phi}_T - \phi_0\right)\right) \\
& + 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H} \hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right), F_{\theta_0}^\intercal \sqrt{T}\left(\hat{\theta}_T - \theta_0\right)\right) \\
& - 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}} - \boldsymbol{\psi}_0\right)\right), F_{\phi_0}^\intercal \sqrt{T}\left(\hat{\phi}_T - \phi_0\right)\right) \\
& - 2\sqrt{\pi} \mathrm{Acov}\left(\sqrt{T}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}} - \boldsymbol{\psi}_0\right)\right), F_{\theta_0}^\intercal \sqrt{T}\left(\hat{\theta}_T - \theta_0\right)\right)
\end{aligned}$$

with $F_{\theta_0}^\intercal = \mathrm{p}\lim_{T,\mathcal{T}\to\infty} \nabla_\theta \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\psi}_{\mathcal{T},h}\left(\phi_0, \theta_0, \lambda_0\right)\right)$,
$F_{\phi_0}^\intercal = \mathrm{p}\lim_{T,\mathcal{T}\to\infty} \nabla_\phi \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\psi}_{\mathcal{T},h}\left(\phi_0, \theta_0, \lambda_0\right)\right)$, $\mathcal{T}/T \to \pi$ *and, finally,* λ_0 *is the minimizer of the moment conditions in Eq. (37) for* $\Delta\sqrt{T} \to 0$, $\Delta T \to \infty$, $\mathcal{T}/T \to \pi \in (0,1)$.

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

3.4 Bootstrap Standard Errors

The limiting covariance matrices in the Propositions 2-4 above, $\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$, are difficult to estimate, as this would require the computation of several numerical derivatives. Also, $\mathbf{V}_3$ reflects the contribution of parameter estimation error. Hence, we do not have a closed form expression for the standard errors. A viable route is then to rely on bootstrap standard errors. Our estimation procedure is based on an hybrid between Indirect Inference and Simulated GMM. Because the auxiliary models are potentially dynamically misspecified, their score is not necessarily a martingale difference sequence. Thus, a natural solution is to use the block bootstrap, which takes into account possible correlation in the score of the auxiliary models.

We shall proceed as follows. We draw b overlapping blocks of length l, with $T = bl$, of

$$X_t = (y_{1,t}, \cdots, y_{1,t-k_1}, y_{2,t}, \cdots, y_{2,t-k_2}, s_t, \cdots, s_{t-k_3}),$$

where k_1, k_2, k_3 depend on the lags we use in the auxiliary models. Hereafter, let

$$X_t^* = (y_{1,t}^*, \cdots, y_{1,t-k_1}^*, y_{2,t}^*, \cdots, y_{2,t-k_2}^*, s_t^*, \cdots, s_{t-k_3}^*)$$

be the set of re-sampled observations.

3.4.1 Bootstrap Standard Errors for ϕ

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 4, 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. On the other hand, as the total 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. Broadly speaking, 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 proper re-centering term is necessary (see e.g. Hall and Horowitz 1996).

Let $\tilde{\varphi}_T^*$ be the bootstrap analog of $\tilde{\varphi}_T$, i.e.

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

where $\tilde{\varphi}_{1,T}^*$ and $\tilde{\varphi}_{2,T}^*$ are the estimated parameters of the auxiliary models computed using re-sampled observations, and $\bar{y}_1^*$, $\bar{y}_2^*$, $\hat{\sigma}_1^*$, $\hat{\sigma}_2^*$ are sample averages and standard deviations of $y_{1,t}^*, y_{2,t}^*$. Define,

$$\hat{\phi}_T^* = \arg\min_{\phi \in \Phi_0} \left\| \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\varphi}_{T,h}^{\Delta}(\phi) - \hat{\varphi}_{T,h}^{\Delta}(\hat{\phi}_T) \right) - (\tilde{\varphi}_T^* - \tilde{\varphi}_T) \right) \right\|^2.$$

We compute B bootstrap estimators $\hat{\phi}_{T,i}^*$, as well as the bootstrap covariance matrix, as follows:

$$\hat{\mathbf{V}}_{\phi_0,T,B}^* = \frac{T}{B} \sum_{i=1}^{B} \left\| \hat{\phi}_{T,i}^* - \frac{1}{B} \sum_{i=1}^{B} \hat{\phi}_{T,i}^* \right\|^2.$$

As shown in Appendix B (Proposition B1), we obtain asymptotically valid bootstrap standard errors from $\left(1 + \frac{1}{H}\right) \hat{\mathbf{V}}_{\phi_0,T,B}^*$.

3.4.2 Bootstrap Standard Errors for θ

The model-based stock price series is simulated out of actual samples of the observable factors, and simulated samples for the unobservable factor. Thus, we need to take into account the contribution of $\mathbf{K}_2$, 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:

$$s_{t,h}^{*\Delta}(\theta) = s_0 + s_1 y_{1,t}^* + s_2 y_{2,t}^* + Z_{t,h}^{*\Delta}(\theta_u), \tag{38}$$

where $Z_{t,h}^{*\Delta}(\theta_u)$ is re-sampled from the simulated unobservable process $Z_{t,h}^{\Delta}(\theta_u)$, and use $s_{t,h}^{*\Delta}(\theta)$ to construct $R_{t,h}^{*\Delta}(\theta)$ and $\mathrm{Vol}_{t,h}^{*\Delta}(\theta)$. Define,

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

where $\tilde{\vartheta}_{1,T}^*, \tilde{\vartheta}_{2,T}^*$ are the estimators of the auxiliary models obtained using re-sampled observations, and $\overline{R}^*, \overline{\mathrm{Vol}}^*$ are the sample mean of $R_t^* = \log(s_t^*/s_{t-1}^*)$ and of $\mathrm{Vol}_t^* = \sqrt{6\pi}\cdot\frac{1}{12}\sum_{i=1}^{12}\left|R_{t+1-i}^*\right|$, with s_t^* being the re-sampled series of the observable stock prices process s_t, and

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

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

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

Compute the bootstrap covariance matrix, as

$$\widehat{\mathbf{V}}_{\theta_0,T,B}^* = T\frac{1}{B}\sum_{i=1}^{B}\left\|\hat{\theta}_{T,i}^* - \frac{1}{B}\sum_{i=1}^{B}\hat{\theta}_{T,i}^*\right\|^2.$$

As shown in Appendix B (Proposition B2), we obtain asymptotically valid bootstrap standard errors from $\left(1+\frac{1}{H}\right)\hat{\mathbf{V}}_{\theta_0,T,B}^*$.

3.4.3 Bootstrap Standard Errors for λ

As mentioned already, 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 $l_\mathcal{T}$ and number of blocks $b_\mathcal{T}$, so that $l_\mathcal{T} b_\mathcal{T} = \mathcal{T}$. Also, we need to re-sample the unobservable factor $Z_{t,h}^{\Delta}(\hat{\theta}_T)$ from a sample of length $\mathcal{T}$. Let $\text{VIX}_{t,h}^{*\Delta}\left(\boldsymbol{y}_t^*;\hat{\phi}_T^*,\hat{\theta}_T^*,\lambda\right)$ be the model-based VIX index constructed using $y_{1,t}^*$, $y_{2,t}^*$, $Z_{t,h}^{*\Delta}(\hat{\theta}_T)$ and the bootstrap estimators $\hat{\phi}_T^*$ and $\hat{\theta}_T^*$. Finally, let

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

where $\tilde{\psi}_{1,\mathcal{T}}^*$ 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 $\overline{\text{VIX}}^*, \hat{\sigma}_{\text{VIX}}^*$ are the sample mean and standard deviation of VIX_t^*, and:

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

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

$$\hat{\lambda}_{\mathcal{T}}^* = \arg\min_{\lambda\in\Lambda_0}\left\|\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\psi}_{\mathcal{T},h}^{*\Delta}(\hat{\phi}_T^*,\hat{\theta}_T^*,\lambda) - \hat{\psi}_{\mathcal{T},h}^{\Delta}(\hat{\phi}_T,\hat{\theta}_T,\hat{\lambda}_{\mathcal{T}})\right) - \left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}^* - \tilde{\boldsymbol{\psi}}_{\mathcal{T}}\right)\right)\right\|^2.$$

Construct the bootstrap covariance matrix, as

$$\hat{\mathbf{V}}_{\lambda_0,\mathcal{T},B}^* = \frac{\mathcal{T}}{B}\sum_{i=1}^{B}\left\|\hat{\lambda}_{\mathcal{T},i}^* - \frac{1}{B}\sum_{i=1}^{B}\hat{\lambda}_{\mathcal{T},i}^*\right\|^2.$$

As shown in Appendix B (Proposition B3), we obtain asymptotically valid bootstrap standard errors from $\left(1+\frac{1}{H}\right)\hat{\mathbf{V}}_{\lambda_0,\mathcal{T},B}^*$.

4 Empirical analysis

4.1 Data

Our security data include the S&P Compounded index and the VIX index, as published by the Chicago Board of Exchange. We compute the real stock price as the ratio between the S&P index and the consumer price index, described in detail below. Data for the VIX index are available daily, but only for the period following January 1990. Information related to the CPI and the IP is made available to the market between the 19th and the 23th of every month. To possibly avoid overreaction to releases of information, we sample the S&P Compounded index and the VIX index every 25th of the month.

Our macroeconomic variables include the consumer price index and the index of industrial production for the US, observed monthly from January 1950 to December 2006, for a total of

672 observations. We take these two series to compute the two macroeconomic factors, the gross inflation and the gross industrial production growth, both at a yearly level,

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

where CPI_t is the consumer price index and IP_t is the seasonally adjusted industrial production index, as of month t. As we explained in the Introduction, many theories lead us to expect that stock prices are indeed related to variables tracking the business cycle conditions (see, e.g., Cochrane (2005)), such as the CPI and the IP growth. In fact, we might have used additional variables to model how the pricing kernel relates to the business cycle. At the same time, our objective is to keep the analysis, and variable selection, as simple as possible. Our empirical results can certainly be improved through a more thorough variable selection, although we leave this issue to further investigation.

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 occurring 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 economic 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, stock volatility and volatility risk-premia.

Finally, Figure 1 shows that gross industrial production growth is also procyclical, and over the entire sampling period. Although its volatility drops during the Great Moderation, it is 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.

4.2 Estimation results

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 section 3.4. 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, κ_2 and $\bar{\kappa}_2$. Note, also, that the sign and value of these parameters are those we need to match the impulse-response functions for $y_{1,t}$ and $y_{2,t}$ that we see in the data (not reported here, for space reasons). These feedback effects do have asset pricing implications, as we shall see 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 useful property, from an asset pricing perspective.

Table 2 reports parameter estimates and standard errors for (i) the parameters relating the two macroeconomic factors, $y_{1,t}$ and $y_{2,t}$, and the unobservable factor, $y_{3,t}$, to the real stock price, 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 outlined in section 3.4. 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. Note, the literature on long run risks started by Bansal and Yaron (2004) emphasizes the asset pricing importance of long run risks affecting the expected consumption growth rate. Interestingly, the presence of a very persistent factor affecting the stock returns and volatility dynamics emerges quite nitidly from our estimation.

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. The 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 point in time t (say), the unobservable factor is set equal to its average across 1000 simulations at time t, $\bar{y}_{3,t} = \frac{1}{1000}\sum_{i=1}^{1000} y_{3,t}^i$ (where $y_{3,t}^i$ is the value of $y_{3,t}$ at simulation i), and the stock price is computed through Eq. (28), using all the estimated parameters. Given the simulated stock price, we compute returns (displayed in the top panel) and volatility (displayed in the bottom panel).

The model appears to capture the procyclical nature of stock returns and the countercyclical behavior of volatility. It generates *all* the stock market drops occurred during the NBER recessions, and *all* the volatility swings occurred during the NBER recessions, including the dramatic spike of the 1975 recession. In the data, average 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 these figures can be attributable to the variation of the macroeconomic factors? After all, the key innovation of our model is the introduction of these factors, on top of a standard unobservable factor. Naturally, the cyclical properties of stock returns and volatility that we see in Figure 2 can only be 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 volatility and its standard deviation. The following table reports the results. 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 quite marginal role as a determinant of stock market volatility, and its cyclical properties.

Percentage reductions in stock volatility and vol of vol

	average	std dev
without y_1	≈ 0	≈ 0
without y_2	11%	78%
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 $y_{3,t}$ is taken 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 observe in the data. When $y_{2,t}$ is frozen, the standard deviation of the model-implied stock volatility drops dramatically to 78%.

When, instead, $y_{3,t}$ is frozen, we even observe an increase in the variability of stock volatility, of about 95%. This last 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 asset returns. Indeed, Figure 3 shows that the level of stock volatility is quite high until the recession occuring 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 in the case in which 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 neat conclusion: the level of stock volatility can not be explained by macroeconomic variables only. Instead, some unobservable factor is needed. At the same time, the same unobservable factor can not explain the variability in stock volatility. Our empirical results suggest that the volatility of stock volatility, can be explained by the cylical variations in stock volatility, which our model captures through the relation between asset returns and industrial production growth. In turns, these swings are amplified by the presence of the unobservable factor.

Table 3 reports parameter estimates and standard errors for the vector of the risk-premia coefficients $\lambda^{\intercal} = \left(\lambda_{1(1)}, \lambda_{2(1)}, \lambda_{1(2)}, \lambda_{2(2)}, \lambda_{1(3)}, \lambda_{2(3)}\right)$, in the risk-premium process in Eq. (19). The estimates, which are all significant, are obtained through the third, and final step, of our estimation procedure, described in section 3.3. Standard errors are computed through the block-bootstrap outlined in section 3.4.

Our estimates imply that the risk-premia processes are all positive and countercyclical, as the sign of the estimated values for both the loadings of gross inflation, $\left(\beta_1\lambda_{1(1)} + \lambda_{2(1)}\right)$, and industrial production, $\left(\beta_2\lambda_{1(2)} + \lambda_{2(2)}\right)$, to the risk-premium process in Eq. (19), is negative. So in bad times, the risk-premium 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 4 (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 4 also shows the dynamics of expected future volatility, under the physical measure. This expected volatility is certainly countercyclical, although it does not display the large variations the model predicts for its risk-neutral counterparts, the VIX index. The VIX index predicted by the model is countercyclical because the risk-premia 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 countercylical, under the physical probability, only because of the third effect. Figure 5 reveals the "tilt" in the future paths of industrial production growth that we need, in order to make our model match the data. 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.

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. In bad times, when gross inflation is lower than its long-run mean, μ_2, future inflation is expected to lower even more, under the risk-neutral probability. But the feedback parameter, $\bar{\kappa}_2$, is positive and large in value, which makes the risk-neutral expectation of industrial production worsen even more. Therefore, although our previous findings suggest 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 4 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 of 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-premia 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 6 provides scatterplots of the volatility risk-premium against inflation and industrial production. The top panel reveals that volatility risk-premium does not display a neat relation with inflation. Instead, the bottom panel reveals a neat and negative relation between the volatility risk-premium and industrial production growth, and suggests the presence of two "regimes", one regime occurring during the 1991 recession, and the other, more severe, occurring during the 2001 recession.

5 Conclusion

This paper develops a model that analyzes how stock market volatility and volatility risk-premia relate to the development of the business cycle. The model's assumption is that the price is uniquely a function of two macroeconomic factors, inflation and industrial production growth, and one unobservable factor. The relation between the asset price, stock volatility and volatility risk-premia are consistent with the assumption of no-arbitrage. This key aspect differentiates our approach from previous models with stochastic volatility, in which volatility was specified exogenously to the price process. The second important feature of our no-arbitrage framework is that we model volatility by making explicit reference to macroeconomic data. The current stochastic volatility literature, instead, is still dealing with models in which volatility is driven by some unobservable factors, just as in the initial seminal contributions of Hull and White (1987) and Heston (1993).

In taking our model to data, we face a delicate identification issue, arising because volatility is endogenous in our model: the same variables driving the payoff process and the volatility of the pricing kernel, and hence, the asset price, are those that drive volatility and volatility-related risk-premia. We implement a three-step procedure in which we estimate, sequentially, (i) the parameters of the macroeconomic factors, (ii) those linking stock returns and volatility to the macroeconomic factors and the unobservable factor and, (iii) those related to risk-aversion correction, and use the new VIX index to identify the parameters related to risk-aversion correction. Our empirical results suggest that the level of stock market volatility we have experienced since the 1950s is largely attributable to the presence of some volatile and quite persistent unobservable factor. This unobservable factor, alone, can not explain the rich dynamics of volatility - only (and partially) its level. We show, instead, that industrial production is largely responsible for the random fluctuations of stock market volatility around its level, and that inflation plays, instead, a quite limited role. Our model predicts the same countercyclical properties of volatility that we see in the data and the large spikes occurring over severe recessions. It also suggests that volatility risk-premia are countercyclical, certainly more so than stock volatility alone. We find that such a countercyclical behavior arises because the risk-compensation for the fluctuation in the macroeconomic factors is large and countercyclical, and that these aspects might make the volatility premium anticipate the development of the business cycle in bad times.

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

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

Then, by Duffie and Kan (1996) (unnumbered theorem, p. 388), there exists a unique strong solution to Eq. (1) for which $\boldsymbol{V}(\boldsymbol{y}(t))_{(ii)} > 0$ for all t almost surely.

We apply these conditions to the diffusion in Eq. (18). Condition (i) collapses to,

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

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,

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

where

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

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 E denotes the expectation taken under the physical probability P, and $\delta^\xi(t)$ 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. (7) in the main text follows, once we show that $\int_t^\infty E[\delta^\xi(h)]\mathrm{d}h < \infty$.

Next, by Eq. (A4),

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

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 $\boldsymbol{\Lambda}$, Eq. (A6) can be rearranged to yield the following ordinary differential equation,

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

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

Let us 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,

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

The solution to this system is,

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

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

As regards 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,

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

which is the condition in (A2). ■

B. Proofs for Section 3

5.1 B.1. Proofs of Propositions 2, 3 and 4

The compact set Φ in Section 3.1 is defined as:

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

and

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

Furthermore, we let ϕ_0 and θ_0 be the solutions to the two limit problems,

$$\phi_0 = \arg\min_{\phi\in\Phi_0} \operatorname*{plim}_{T\to\infty,\Delta\to 0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi) - \tilde{\varphi}_T \right\|^2,$$

and

$$\theta_0 = \arg\min_{\theta\in\Theta_0} \operatorname*{plim}_{T\to\infty,\Delta\to 0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{\Delta}(\theta) - \tilde{\vartheta}_T \right\|^2.$$

Finally, we define the limit problem for the estimator of the risk-premium parameters,

$$\lambda_0 = \arg\min_{\lambda\in\Lambda} \operatorname*{plim}_{T\to\infty,\Delta\to 0} \left\| \frac{1}{H}\sum_{h=1}^{H} \hat{\psi}_{T,h}^{\Delta}(\hat{\phi}_T, \hat{\theta}_T, \lambda) - \tilde{\boldsymbol{\psi}}_T \right\|^2.$$

We are now ready to prove the propositions in Section 3.

Proof of Proposition 2: By the conditions in (A1) of Appendix A, we have that A1(iii) ensure that the diffusion $(y_1(t), y_2(t))$ is the unique strong solution to Eq. (18), and has positive-definite covariance matrix with probability one. A1(ii) ensure that $(y_1(t), y_2(t))$ is geometrically ergodic and the skeleton $(y_{1,t}, y_{2,t})$ is geometrically β-mixing. For any $\varphi \in \Phi_0$, $y_{1,t,h}(\phi), y_{2,t,h}(\phi)$, the paths simulated with $\Delta = 0$, are geometrically ergodic with positive-definite covariance matrix with probability one. The simulated skeleton, $(y_{1,t,\Delta,h}^{\phi}, y_{2,t,\Delta,h}^{\phi})$, is also geometrically β-mixing and, given (18), is at least twice-continuously differentiable in any open neighborhood of ϕ_0.

We claim that $\hat{\phi}_T - \phi_0 = o_p(1)$, 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 ϕ_0,

$$\begin{aligned}
0 &= \nabla_\phi \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\hat{\phi}_T)\right)^{\intercal} \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\hat{\phi}_T) - \tilde{\varphi}_T\right) \\
&= \nabla_\phi \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\hat{\phi}_T)\right)^{\intercal} \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi_0) - \tilde{\varphi}_T\right) \\
&\quad + \nabla_\phi \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\hat{\phi}_T)\right)^{\intercal} \nabla_\phi \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\bar{\phi}_T)\right) \left(\hat{\phi}_T - \phi_0\right),
\end{aligned}$$

where $\bar{\phi}_T$ is some convex combination of $\hat{\phi}_T$ and ϕ_0. By the uniform law of large numbers, $\sup_{\phi\in\Phi_0} \left|\nabla_\phi \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi)\right) - \mathbf{D}_1(\phi)\right| = o_p(1)$, and as $\hat{\phi}_T - \phi_0 = o_p(1)$, $\nabla_\phi \left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\bar{\phi}_T)\right) - \mathbf{D}_1 = o_p(1)$. Hence,

$$\sqrt{T}\left(\hat{\phi}_T - \phi_0\right) = -\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1} \mathbf{D}_1^{\intercal} \left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi_0) - \varphi_0\right) - \sqrt{T}\left(\tilde{\varphi}_T - \varphi_0\right)\right) + o_p(1).$$

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

$$\sqrt{T}\left(\hat{\varphi}_{T,h}^{\Delta}(\phi) - \hat{\varphi}_{T,h}(\phi_0)\right) = o_p(1).$$

Let $Y_{t,h}^{\Delta}(\phi_0) = \left(y_{1,t-j,\Delta,h}^{\phi}, y_{1,t-j,\Delta,h}^{\phi},\ j \in \{12, 24\}\right)$, then

$$\begin{aligned}
&\sqrt{T}\left(\hat{\varphi}_{1,T,h}^{\Delta}(\phi_0) - \hat{\varphi}_{1,T,h}(\phi_0)\right) \\
= &\left(\frac{1}{T}\sum_{t=25}^{T} \mathbf{Y}_{t,h}^{\phi_0} \mathbf{Y}_{t,h}^{\phi_0 \intercal}\right)^{-1} \sqrt{T}\left(\frac{1}{T}\sum_{t=25}^{T}\left(\mathbf{Y}_{t,h}^{\Delta,\phi_0} y_{1,t,h}^{\Delta,\phi_0} - \mathbf{Y}_{t,h}^{\phi_0} y_{1,t,h}^{\phi_0}\right)\right) \\
&+\sqrt{T}\left(\left(\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_{t,h}^{\Delta,\phi_0}\mathbf{Y}_{t,h}^{\Delta,\phi_0\intercal}\right)^{-1} - \left(\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_{t,h}^{\phi_0}\mathbf{Y}_{t,h}^{\phi_0\intercal}\right)^{-1}\right)\left(\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_{t,h}^{\Delta,\phi_0} y_{1,t,h}^{\Delta,\phi_0}\right) \qquad \text{(B1)}
\end{aligned}$$

As for the first term on the RHS of (B1), $\left(\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_{t,h}^{\phi_0}\mathbf{Y}_{t,h}^{\phi_0\intercal}\right)^{-1} = O_p(1)$ and by Theorem 2.3 in Pardoux and Talay (1985), we have, for $\varepsilon > 0$ and $\sqrt{T}\Delta \to 0$,

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

The second term on the right hand side of Eq. (B1) can be dealt with similarly. Thus, we have:

$$\begin{aligned}
&\mathrm{Avar}\left(\sqrt{T}\left(\hat{\phi}_T - \phi_0\right)\right) \\
= &\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1}\mathbf{D}_1^{\intercal}\mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}(\phi_0) - \varphi_0\right) - \sqrt{T}\left(\tilde{\varphi}_T - \varphi_0\right)\right)\mathbf{D}_1\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1},
\end{aligned}$$

where,

$$\begin{aligned}
&\mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}(\phi_0) - \varphi_0\right) - \sqrt{T}\left(\tilde{\varphi}_T - \varphi_0\right)\right) \\
&= \mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}(\phi_0) - \varphi_0\right)\right) + \mathrm{Avar}\left(\sqrt{T}\left(\tilde{\varphi}_T - \varphi_0\right)\right) \\
&- 2\mathrm{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}(\phi_0) - \varphi_0\right), \sqrt{T}\left(\tilde{\varphi}_T - \varphi_0\right)\right).
\end{aligned}$$

As the simulated paths are independent of the sample paths,

$$\mathrm{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}(\phi_0) - \varphi_0\right), \sqrt{T}\left(\tilde{\varphi}_T - \varphi_0\right)\right) = 0.$$

As simulated paths are identically distributed and independent across different simulation replications,

$$
\begin{aligned}
&\text{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}\left(\phi_0\right)-\varphi_0\right)\right)\\
&=\frac{1}{H^2}\sum_{h=1}^{H}\text{Avar}\left(\sqrt{T}\left(\hat{\varphi}_{T,h}\left(\phi_0\right)-\varphi_0\right)\right)+\frac{1}{H^2}\sum_{h=1}^{H}\sum_{h'\neq h}^{H}\text{Acov}\left(\sqrt{T}\left(\hat{\varphi}_{T,h}\left(\phi_0\right)-\varphi_0\right),\sqrt{T}\left(\hat{\varphi}_{T,h'}\left(\phi_0\right)-\varphi_0\right)\right)\\
&=\frac{1}{H}\text{Avar}\left(\sqrt{T}\left(\hat{\varphi}_{T,1}\left(\phi_0\right)-\varphi_0\right)\right)
\end{aligned}
$$

Finally, given A1, $\text{Avar}\left(\sqrt{T}\left(\tilde{\varphi}_{T,1}-\varphi_0\right)\right)=\text{Avar}\left(\sqrt{T}\sqrt{T}\left(\tilde{\varphi}_{T}-\varphi_0\right)\right)=\mathbf{J}_1$, and so

$$\text{Avar}\left(\sqrt{T}\left(\hat{\phi}_T-\phi_0\right)\right)=\left(1+\frac{1}{H}\right)\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1}\mathbf{D}_1^{\intercal}\mathbf{J}_1\mathbf{D}_1\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1}.$$

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

The expressions for σ^2, $\nabla_y\sigma^2$ and $\nabla_{yy}\sigma^2$, $\boldsymbol{c}$, $\boldsymbol{D}$ in Eq. (35). Under the conditions in Section 2, we have that in Eq. (35),

$$
\begin{aligned}
\sigma^2\left(\boldsymbol{y}\right) &= \frac{\sum_{j=1}^{3}s_j^2\left(\alpha_j+\beta_j y_j\right)}{s\left(\boldsymbol{y}\right)^2}\\
\nabla_{y_j}\sigma^2\left(\boldsymbol{y}\right) &= \frac{s_j^2\beta_j-2\sigma^2\left(\boldsymbol{y}\right)s\left(\boldsymbol{y}\right)s_j}{s\left(\boldsymbol{y}\right)^2}\\
\nabla_{y_jy_j}\sigma^2\left(\boldsymbol{y}\right) &= -2\frac{s_j}{s\left(\boldsymbol{y}\right)^2}\left(\frac{s_j^2\beta_j}{s\left(\boldsymbol{y}\right)}+s\left(\boldsymbol{y}\right)\nabla_{y_j}\sigma^2\left(\boldsymbol{y}\right)-s_j\sigma^2\left(\boldsymbol{y}\right)\right)
\end{aligned}
$$

and the coefficients $\boldsymbol{c}$ and $\boldsymbol{D}$ in Eqs. (10)-(11) are given by:

$$
\begin{aligned}
\boldsymbol{c} &= \begin{bmatrix} \kappa_1\mu_1+\bar{\kappa}_1\mu_2-\alpha_1\lambda_{1(1)} \\ \bar{\kappa}_2\mu_1+\kappa_2\mu_2-\alpha_2\lambda_{1(2)} \\ \kappa_3\mu_3-\alpha_3\lambda_{1(3)} \end{bmatrix}\\
\boldsymbol{D} &= \begin{bmatrix} \kappa_1+\lambda_{1(1)}\beta_1+\lambda_{2(1)} & \bar{\kappa}_1 & 0 \\ \bar{\kappa}_2 & \kappa_2+\lambda_{1(2)}\beta_2+\lambda_{2(2)} & 0 \\ 0 & 0 & \kappa_3+\lambda_{1(3)}\beta_3+\lambda_{2(3)} \end{bmatrix}.
\end{aligned}
$$

with $\mu_3\equiv 1$ for identifiability.

Proof of Proposition 3: By the same argument utilized in the proof of Proposition 2,

$$\sqrt{T}\left(\hat{\theta}_T-\theta_0\right)=-\left(\mathbf{D}_2^{\intercal}\mathbf{D}_2\right)^{-1}\mathbf{D}_2^{\intercal}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}^{\Delta}\left(\theta_0\right)-\vartheta_0\right)-\sqrt{T}\left(\tilde{\vartheta}_T-\vartheta_0\right)\right)+o_p(1).$$

Thus,

$$\text{Avar}\left(\sqrt{T}\left(\hat{\theta}_T-\theta_0\right)\right)=\left(\mathbf{D}_2^{\intercal}\mathbf{D}_2\right)^{-1}\mathbf{D}_2^{\intercal}\cdot\mathbf{J}_0\cdot\mathbf{D}_2\left(\mathbf{D}_2^{\intercal}\mathbf{D}_2\right)^{-1},$$

where

$$\mathbf{J}_0 = \text{Avar}\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}^{\Delta}(\theta_0) - \vartheta_0\right) + \text{Avar}\sqrt{T}\left(\tilde{\vartheta}_T - \vartheta_0\right)$$
$$- 2\text{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}^{\Delta}(\theta_0) - \vartheta_0\right), \sqrt{T}\left(\tilde{\vartheta}_T - \vartheta_0\right)\right)$$

Let $\hat{\vartheta}_{T,h}(\theta_0)$ be the the unfeasible estimator, obtained by simulating continuous paths for the unobservable factor $Z(t)$, i.e. $Z_{t,h}(\theta)$. By the same argument as that in the proof of Proposition 2,

$$\text{Avar}\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}^{\Delta}(\theta_0) - \vartheta_0\right) = \text{Avar}\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}(\theta_0) - \vartheta_0\right).$$

In the current context, 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 simulation replications, and are no longer independent of the actual sample paths of stock price and volatility. Thus,

$$\begin{aligned}
&\text{Avar}\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}(\theta_0) - \vartheta_0\right)\\
&= \frac{1}{H}\text{Avar}\left(\sqrt{T}\left(\hat{\vartheta}_{T,1}(\theta_0) - \vartheta_0\right)\right) + \frac{H(H-1)}{H^2}\text{Acov}\left(\sqrt{T}\left(\hat{\vartheta}_{T,1}(\theta_0) - \vartheta_0\right), \sqrt{T}\left(\hat{\vartheta}_{T,h}(\theta_0) - \vartheta_0\right)\right)\\
&= \frac{1}{H}\mathbf{J}_2 + \frac{H(H-1)}{H^2}\mathbf{K}_2,
\end{aligned}$$

and

$$\begin{aligned}
\text{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}(\theta_0) - \vartheta_0\right), \sqrt{T}\left(\tilde{\vartheta}_T - \vartheta_0\right)\right) &= \frac{1}{H}\sum_{h=1}^{H}\text{Acov}\left(\sqrt{T}\left(\hat{\vartheta}_{T,h}(\theta_0) - \vartheta_0\right), \sqrt{T}\left(\tilde{\vartheta}_T - \vartheta_0\right)\right)\\
&= \mathbf{K}_2.
\end{aligned}$$

Finally, because $\text{Avar}\sqrt{T}\left(\tilde{\vartheta}_T - \vartheta_0\right) = \text{Avar}\sqrt{T}\left(\hat{\vartheta}_{T,1}(\theta_0) - \vartheta_0\right) = \mathbf{J}_2$, it follows that

$$\text{Avar}\left(\sqrt{T}\left(\hat{\theta}_T - \theta_0\right)\right) = \left(1 + \frac{1}{H}\right)\left(\mathbf{D}_2^\intercal\mathbf{D}_2\right)^{-1}\mathbf{D}_2^\intercal\left(\mathbf{J}_2 - \mathbf{K}_2\right)\mathbf{D}_2\left(\mathbf{D}_2^\intercal\mathbf{D}_2\right)^{-1}.$$

Proof of Proposition 4: Given the first order conditions, and by a mean value expansion around λ_0,

$$\begin{aligned}
&\sqrt{\mathcal{T}}\left(\hat{\lambda}_{\mathcal{T}} - \lambda_0\right)\\
&= -\left(\nabla_\lambda\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\mathcal{T},h}^{\Delta}(\hat{\phi}_T, \hat{\theta}_T, \hat{\lambda}_{\mathcal{T}})\right)^\intercal \nabla_\lambda\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\mathcal{T},h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \bar{\lambda}_{\mathcal{T}}\right)\right)\right)^{-1}\\
&\nabla_\lambda\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\mathcal{T},h}^{\Delta}(\hat{\phi}_T, \hat{\theta}_T, \hat{\lambda}_{\mathcal{T}})\right)^\intercal \sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\mathcal{T},h}^{\Delta}(\hat{\phi}_T, \hat{\theta}_T, \lambda_0) - \tilde{\boldsymbol{\psi}}_{\mathcal{T}}\right),
\end{aligned}$$

where $\bar{\lambda}_{\mathcal{T}}$ is some convex combination of $\hat{\lambda}_{\mathcal{T}}$ and λ_0. Given Proposition 2 and Proposition 3, by the uniform law of large numbers $\hat{\lambda}_{\mathcal{T}} - \lambda_0 = o_p(1)$ and $\sup_{\lambda\in\Lambda,\theta\in\Theta,\psi\in\Psi}\left|\nabla_\lambda\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\mathcal{T},h}^{\Delta}(\phi,\theta,\lambda)\right) - \mathbf{D}_3(\phi,\theta,\lambda)\right| =$

$o_p(1)$. Therefore, $\nabla_\lambda \left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}(\hat{\phi}_T,\hat{\theta}_T,\hat{\lambda}_T)\right) - \mathbf{D}_3 = o_p(1)$, and, hence,

$$\sqrt{\mathcal{T}}\left(\hat{\lambda}_{\mathcal{T}}-\lambda_0\right) = -\left(\mathbf{D}_3^{\intercal}\mathbf{D}_3\right)^{-1}\mathbf{D}_3^{\intercal}\left(\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}(\hat{\phi}_T,\hat{\theta}_T,\lambda_0)-\boldsymbol{\psi}_0\right)-\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right)+o_p(1).$$

We have,

$$\begin{aligned}
&\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}(\hat{\phi}_T,\hat{\theta}_T,\lambda_0)-\boldsymbol{\psi}_0\right)\\
&=\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right)+\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\psi}^{\Delta}_{\mathcal{T},h}(\hat{\phi}_T,\hat{\theta}_T,\lambda_0)-\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\hat{\phi}_T,\theta_0,\lambda_0\right)\right)\right)\\
&+\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\hat{\phi}_T,\theta_0,\lambda_0\right)-\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)\right)\right)\\
&=\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right)+\nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\hat{\phi}_T,\bar{\theta}_T,\lambda_0\right)\right)\sqrt{\mathcal{T}}\left(\hat{\theta}_T-\theta_0\right)\\
&+\nabla_\phi\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\bar{\phi}_T,\theta_0,\lambda_0\right)\right)\sqrt{\mathcal{T}}\left(\hat{\phi}_T-\phi_0\right)\\
&=\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right)+F^{\intercal}_{\theta_0}\sqrt{\pi}\sqrt{T}\left(\hat{\theta}_T-\theta_0\right)+F^{\intercal}_{\phi_0}\sqrt{\pi}\sqrt{T}\left(\hat{\phi}_T-\phi_0\right)+o_p(1),
\end{aligned}$$

where $\pi = \lim_{T,\mathcal{T}\to\infty}\mathcal{T}/T$, $\bar{\theta}_T$ is some convex combination of $\hat{\theta}_T$ and θ_0, $\bar{\phi}_T$ is some convex combination of $\hat{\phi}_T$ and ϕ, and:

$$F^{\intercal}_{\theta_0}=\underset{T,\mathcal{T}\to\infty}{\text{plim}}\,\nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}(\hat{\phi}_T,\bar{\theta}_T,\lambda_0)\right)\text{ and }F^{\intercal}_{\phi_0}=\underset{T,\mathcal{T}\to\infty}{\text{plim}}\,\nabla_\phi\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_T,\theta_0,\lambda_0\right)\right).$$

We have:

$$\begin{aligned}
&\text{Avar}\sqrt{\mathcal{T}}\left(\hat{\lambda}_{\mathcal{T}}-\lambda_0\right)\\
=\;&\left(\mathbf{D}_3^{\intercal}\mathbf{D}_3\right)^{-1}\mathbf{D}_3^{\intercal}\left(\text{Avar}\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right)+\text{Avar}\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right.\\
&-2\text{Acov}\left(\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right),\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right)+\pi F^{\intercal}_{\theta_0}\text{Avar}\left(\sqrt{T}\left(\hat{\theta}_T-\theta_0\right)\right)F_{\theta_0}\\
&+\pi F^{\intercal}_{\phi_0}\text{Avar}\left(\sqrt{T}\left(\hat{\phi}_T-\phi_0\right)\right)F_{\phi_0}+2\pi\text{Acov}\left(F^{\intercal}_{\phi_0}\sqrt{T}\left(\hat{\phi}_T-\phi_0\right),F^{\intercal}_{\theta_0}\sqrt{T}\left(\hat{\theta}_T-\theta_0\right)\right)\\
&+2\sqrt{\pi}\text{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right),F^{\intercal}_{\phi_0}\sqrt{T}\left(\hat{\phi}_T-\phi_0\right)\right)\\
&+2\sqrt{\pi}\text{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right),F^{\intercal}_{\theta_0}\sqrt{T}\left(\hat{\theta}_T-\theta_0\right)\right)\\
&-2\sqrt{\pi}\text{Acov}\left(\sqrt{T}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right),F^{\intercal}_{\phi_0}\sqrt{T}\left(\hat{\phi}_T-\phi_0\right)\right)\\
&\left.-2\sqrt{\pi}\text{Acov}\left(\sqrt{T}\left(\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right),F^{\intercal}_{\theta_0}\sqrt{T}\left(\hat{\theta}_T-\theta_0\right)\right)\right)\mathbf{D}_3\left(\mathbf{D}_3^{\intercal}\mathbf{D}_3\right)
\end{aligned}\qquad\text{(B2)}$$

Let $\hat{\psi}_{\mathcal{T},h}(\hat{\phi}_T, \hat{\theta}_T, \lambda_0)$ be the estimator obtained in the case we computed the model-based VIX using continuous simulated path for the unobservable factor $Z_{t,h}(\theta)$. By a similar argument as in Proposition 2,

$$\text{Avar}\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}^{\Delta}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right)=\text{Avar}\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right)$$

By the same argument as in Proposition 3,

$$\begin{aligned}
&\text{Avar}\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right)+\text{Avar}\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\\
&-2\text{Acov}\left(\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\mathcal{T},h}\left(\phi_0,\theta_0,\lambda_0\right)-\boldsymbol{\psi}_0\right),\sqrt{\mathcal{T}}\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}_0\right)\right)\\
=\;&\left(1+\frac{1}{H}\right)\left(\mathbf{J}_3-\mathbf{K}_3\right).
\end{aligned}$$

Hence, given Propositions 2 and 3,

$$\begin{aligned}
&\text{Avar}\sqrt{\mathcal{T}}\left(\hat{\lambda}_{\mathcal{T}}-\lambda_0\right)\\
=\;&\left(\mathbf{D}_3^{\intercal}\mathbf{D}_3\right)^{-1}\mathbf{D}_3^{\intercal}\left(\left(1+\frac{1}{H}\right)\left(\left(\mathbf{J}_3-\mathbf{K}_3\right)+\pi F_{\theta_0}^{\intercal}\left(\mathbf{D}_2^{\intercal}\mathbf{D}_2\right)^{-1}\mathbf{D}_2^{\intercal}\left(\mathbf{J}_2-\mathbf{K}_2\right)\mathbf{D}_2\left(\mathbf{D}_2^{\intercal}\mathbf{D}_2\right)^{-1}F_{\theta_0}\right.\right.\\
&\left.+\pi F_{\phi_0}^{\intercal}\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1}\mathbf{D}_1^{\intercal}\mathbf{J}_1\mathbf{D}_1\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1}F_{\phi_0}\right)+2\pi C_{\theta,\phi}\\
&\left.+2\sqrt{\pi}\left(C_{\psi_h,\phi}+C_{\psi_h,\theta}-C_{\psi,\phi}+C_{\psi,\theta}\right)\right)\mathbf{D}_3\left(\mathbf{D}_3^{\intercal}\mathbf{D}_3\right)^{-1},
\end{aligned}$$

where $C_{\theta,\phi}, C_{\psi_h,\phi}, C_{\psi_h,\theta}, C_{\psi,\phi}, C_{\psi,\theta}$ denote the last five asymptotic covariance terms on the RHS of (B2)

B.2. Proofs for the bootstrap standard errors

Hereafter, let P^* be the probability measure governing the re-sampled series, and E^*, var^* denote the mean and the variance taken with respect to P^*; further $O_p^*(1)$ and $o_p^*(1)$ denote a term bounded in probability and converging to zero in probability, according to P^*, conditional on the sample and for all samples but a set of probability measure approaching zero.

Proposition B1: *If, as $T, B \to \infty$, $l/T^{1/2} \to 0$, then:*

$$P\left(\omega : P^*\left(\left|\left(1+\frac{1}{H}\right)\widehat{\mathbf{V}}^*_{\phi_0,T,B}-\mathbf{V}_1\right|>\varepsilon\right)\right)\to 0.$$

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

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

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

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

The Proposition follows, once we show that:

$$\mathrm{E}^*\left(\sqrt{T}\left(\tilde{\varphi}^*_T-\tilde{\varphi}_T\right)\right)=o_p(1), \tag{B3}$$

$$\mathrm{var}^*\left(\sqrt{T}\left(\tilde{\varphi}^*_T-\tilde{\varphi}_T\right)\right)=\mathrm{var}\left(\sqrt{T}\left(\tilde{\varphi}_T-\varphi_0\right)\right)+O_p(l/\sqrt{T}), \tag{B4}$$

and for $\delta>0$,

$$\mathrm{E}^*\left(\left(\sqrt{T}\left(\tilde{\varphi}^*_T-\tilde{\varphi}_T\right)\right)^{2+\delta}\right)=O_p(1). \tag{B5}$$

Indeed, under conditions (B3)-(B4), we have that by the uniform law of large numbers, $\left|\nabla_\phi\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}(\hat{\phi}^*_T)\right)-\mathbf{D}_1\right|=o^*_p(1)$. Hence,

$$\sqrt{T}\left(\hat{\phi}^*_T-\hat{\phi}_T\right)=\left(\mathbf{D}_1^{\intercal}\mathbf{D}_1\right)^{-1}\mathbf{D}_1^{\intercal}\sqrt{T}\left(\tilde{\varphi}_T-\tilde{\varphi}^*_T\right)+o^*_p(1).$$

and, given (B4), and recalling that $l/\sqrt{T}\to 0$,

$$\mathrm{var}^*\left(\sqrt{T}\left(\tilde{\varphi}^*_T-\tilde{\varphi}_T\right)\right)=\mathrm{Avar}\left(\sqrt{T}\left(\tilde{\varphi}_T-\varphi_0\right)\right)+o_p(1).$$

Given (B5), the statement follows by Theorem 1 in Goncalves and White (2005).

Let us show (B3),(B4) and (B5). We have,

$$\sqrt{T}\left(\tilde{\varphi}^*_T-\tilde{\varphi}_T\right)=\left(\sqrt{T}\left(\tilde{\varphi}^*_{i,T}-\tilde{\varphi}_{i,T}\right),\sqrt{T}\left(\bar{y}^*_i-\bar{y}_i\right),\sqrt{T}\left(\hat{\sigma}^{*2}_i-\hat{\sigma}^2_i\right),\ i=1,2\right)^{\intercal}.$$

Since each component of $\sqrt{T}\left(\tilde{\varphi}^*_T-\tilde{\varphi}_T\right)$ can be dealt with in the same way, we only consider $\sqrt{T}\left(\tilde{\varphi}^*_{1,T}-\tilde{\varphi}_{1,T}\right)$. By the first order conditions,

$$\begin{aligned}\sqrt{T}\left(\tilde{\varphi}^*_{1,T}-\tilde{\varphi}_{1,T}\right)&=\left(\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}^*_t\mathbf{Y}^{*\intercal}_t\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=25}^{T}\mathbf{Y}^*_t\left(\mathbf{y}^*_{1,t}-\mathbf{Y}^{*\intercal}_t\tilde{\varphi}_{1,T}\right)\\&=\left(\mathrm{E}\left(\mathbf{Y}_t\mathbf{Y}_t^{\intercal}\right)\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=25}^{T}\mathbf{Y}^*_t\left(\mathbf{y}^*_{1,t}-\mathbf{Y}^{*\intercal}_t\tilde{\varphi}_{1,T}\right)+o^*_p(1),\end{aligned}$$

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

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

This proves (B3). Next,

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

where $\widetilde{\epsilon}_{1,t}=y_{1,t}-\mathbf{Y}_t^T\tilde{\varphi}_{1,T}$. This proves (B4). Finally, as $\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_t\mathbf{Y}_t$ is full rank, by the same argument as above,

$$
\mathrm{E}^*\left(\left(\sqrt{T}\left(\tilde{\varphi}^*_T-\tilde{\varphi}_T\right)\right)^{2+\delta}\right)\leq\left(\frac{1}{\sqrt{T}}\sum_{t=25}^{T}\mathbf{Y}_t\left(\mathbf{y}_{1,t}-\mathbf{Y}_t^{\intercal}\tilde{\varphi}_{1,T}\right)\right)^{2+\delta}=O_p(1)
$$

This proves (B5).

Proposition B2: *If, as* $T,B\to\infty$, $l/T^{1/2}\to 0$, *then*

$$
P\left(\omega:P^*\left(\left|\left(1+\frac{1}{H}\right)\widehat{\mathbf{V}}^*_{\phi_0,T,B}-\mathbf{V}_2\right|>\varepsilon\right)\right)\to 0.
$$

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

$$
\begin{aligned}
0 &= \nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}^*_T)\right)^{\intercal}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}^*_T)-\hat{\vartheta}^{\Delta}_{T,h}(\hat{\theta}_T)\right)-\left(\tilde{\vartheta}^*_T-\tilde{\vartheta}_T\right)\right) \\
&= \nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}^*_T)\right)^{\intercal}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}_T)-\hat{\vartheta}^{\Delta}_{T,h}(\hat{\theta}_T)\right)-\left(\tilde{\vartheta}^*_T-\tilde{\vartheta}_T\right)\right) \\
&\quad+\nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}^*_T)\right)^{\intercal}\nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}^{*\Delta}_{T,h}\left(\overline{\theta}^*_T\right)\right)\left(\hat{\theta}^*_T-\hat{\theta}_T\right),
\end{aligned}
$$

where and $\overline{\theta}^*_T\in\left(\hat{\theta}^*_T,\hat{\theta}_T\right)$. Hence,

$$
\begin{aligned}
& \sqrt{T}\left(\hat{\theta}^*_T-\hat{\theta}_T\right) \\
= \; & -\left(\nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}^*_T)\right)^{\intercal}\nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}^{*\Delta}_{T,h}\left(\overline{\theta}^*_T\right)\right)\right)^{-1} \\
& \nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}^*_T)\right)^{\intercal}\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}_T)-\hat{\vartheta}^{\Delta}_{T,h}(\hat{\theta}_T)\right)-\left(\tilde{\vartheta}^*_T-\tilde{\vartheta}_T\right)\right)
\end{aligned}
$$

We need to show that:

$$
\mathrm{E}^*\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}_T)-\hat{\vartheta}^{\Delta}_{T,h}(\hat{\theta}_T)\right)\right)\right)=o_p(1) \tag{B6}
$$

$$\text{var}^*\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}_T)-\hat{\vartheta}^{\Delta}_{T,h}(\hat{\theta}_T)\right)\right)\right)=\text{var}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{\Delta}_{T,h}(\hat{\theta}_T)-\vartheta\left(\theta_0\right)\right)\right)\right)+o_p(1) \quad \text{(B7)}$$

$$\text{E}^*\left(\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{T,h}(\hat{\theta}_T)-\hat{\vartheta}^{\Delta}_{T,h}(\hat{\theta}_T)\right)\right)\right)^{2+\delta}\right)<\infty. \quad \text{(B8)}$$

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

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

We only consider $\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{1,T,h}(\hat{\theta}_T)-\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)$, as the remaining terms can be dealt with in the same manner. By the first order conditions,

$$\begin{aligned}
&\text{E}^*\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{1,T,h}(\hat{\theta}_T)-\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\right)\\
&=\frac{1}{H}\sum_{h=1}^{H}\text{E}^*\left(\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_{t,h}(\hat{\theta}_T)\mathbf{Y}^{*\intercal\Delta}_{t}(\hat{\theta}_T)\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_{t}(\hat{\theta}_T)\left(R^{*\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{*\intercal\Delta}_{t}(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\\
&=\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}^{\Delta}_{t,h}(\hat{\theta}_T)\mathbf{Y}^{\intercal\Delta}_{t,h}(\hat{\theta}_T)\right)^{-1}\text{E}^*\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_{t}(\hat{\theta}_T)\left(R^{*\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{*\intercal\Delta}_{t}(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\\
&+\frac{1}{H}\sum_{h=1}^{H}\text{E}^*\left(\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_{t,h}(\hat{\theta}_T)\mathbf{Y}^{*\intercal\Delta}_{t}(\hat{\theta}_T)\right)^{-1}-\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}^{\Delta}_{t,h}(\hat{\theta}_T)\mathbf{Y}^{\intercal\Delta}_{t,h}(\hat{\theta}_T)\right)^{-1}\right)\\
&\times\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_{t}(\hat{\theta}_T)\left(R^{*\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{*\intercal\Delta}_{t}(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\Bigg)\\
&=\text{E}^*\left(I^*_{T,h}\right)+\text{E}^*\left(II^*_{T,h}\right).
\end{aligned}$$

We have,

$$\begin{aligned}
\text{E}^*\left(I^*_{T,h}\right)&=\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}^{\Delta}_{t,h}(\hat{\theta}_T)\mathbf{Y}^{\intercal\Delta}_{t,h}(\hat{\theta}_T)\right)^{-1}\text{E}^*\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_{t}(\hat{\theta}_T)\left(R^{*\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{*\intercal\Delta}_{t}(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\\
&=O_p\left(l/\sqrt{T}\right)\\
&=o_p(1),
\end{aligned}$$

and as $II^*_{T,h}$ is of smaller order that $I^*_{T,h}$, $\mathrm{E}^*(II^*_{T,h}) = o_p(1)$. This proves (B6). As for $h = 1, \cdots, H$

$$\begin{aligned}
&\mathrm{E}^*\left(\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_{t,h}(\hat{\theta}_T)\mathbf{Y}^{*\intercal\Delta}_t(\hat{\theta}_T)\right)^{-1}\right)\\
&=\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}^{\Delta}_{t,h}(\hat{\theta}_T)\mathbf{Y}^{\intercal\Delta}_t(\hat{\theta}_T)\right)^{-1}+o(1)\\
&=\mathrm{E}\left(\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}^{\Delta}_{t,h}(\theta_0)\,\mathbf{Y}^{\intercal\Delta}_t(\theta_0)\right)^{-1}\right)+o_p(1),
\end{aligned}$$

it suffices to show that

$$\begin{aligned}
&\mathrm{var}^*\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_t(\hat{\theta}_T)\left(R^{*\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{*\intercal\Delta}_t(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\right)\\
=\;&\mathrm{Avar}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{\Delta}_t(\hat{\theta}_T)\left(R^{\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{\intercal\Delta}_t(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\right)+o_p(1).
\end{aligned}$$

Because the blocks are all independent,

$$\begin{aligned}
&\mathrm{var}^*\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{*\Delta}_t(\hat{\theta}_T)\left(R^{*\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{*\intercal\Delta}_t(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\right)\\
&=\frac{1}{T}\frac{1}{H^2}\sum_{h=1}^{H}\sum_{h'=1}^{H}\sum_{t=13}^{T}\sum_{s=13}^{T}\mathrm{E}^*\left(\mathbf{Y}^{*\Delta}_t(\hat{\theta}_T)\left(R^{*\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{*\intercal\Delta}_t(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\\
&=\frac{1}{T}\frac{1}{H^2}\sum_{h=1}^{H}\sum_{h'=1}^{H}\sum_{j=13-l}^{l}\sum_{s=13+l}^{T-l}\hat{\epsilon}^{\Delta}_{t,h}\hat{\epsilon}^{\Delta}_{t+j,h}\mathbf{Y}^{\Delta}_t(\hat{\theta}_T)\mathbf{Y}^{\Delta}_{t+j}(\hat{\theta}_T)+o_p(1)\\
&=\mathrm{Avar}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{\Delta}_t(\theta_0)\left(R^{\Delta}_{t,h}(\theta_0)-\mathbf{Y}^{\intercal\Delta}_t(\theta_0)\,\vartheta_{1,0}(\theta_0)\right)\right)\right)+o_p(1),
\end{aligned}$$

where $\hat{\epsilon}^{\Delta}_{t,h} = \mathbf{Y}^{\Delta}_t(\hat{\theta}_T)\left(R^{\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{\intercal\Delta}_t(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)$. This proves (B7). Finally, under the parameter restrictions of Appendix A,

$$\begin{aligned}
&\mathrm{E}^*\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\vartheta}^{*\Delta}_{1,T,h}(\hat{\theta}_T)-\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)^{2+\delta}\right)\\
&=\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}^{\Delta}_t(\hat{\theta}_T)\left(R^{\Delta}_{t,h}(\hat{\theta}_T)-\mathbf{Y}^{\intercal\Delta}_t(\hat{\theta}_T)\hat{\vartheta}^{\Delta}_{1,T,h}(\hat{\theta}_T)\right)\right)\right)^{2+\delta}=O_p(1).
\end{aligned}$$

Proposition B3: *If, as $T, \mathcal{T}, B \to \infty$, $l_{\mathcal{T}}/T^{1/2}_{\mathcal{T}} \to 0$, then*

$$P\left(\omega : P^*\left(\left|\left(1+\frac{1}{H}\right)\hat{\mathbf{V}}^*_{\lambda_0,\mathcal{T},B}-\mathbf{V}_3\right|>\varepsilon\right)\right)\to 0.$$

Proof: By the first order conditions and a mean value expansion around $\hat{\lambda}_{\mathcal{T}}$,

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

where $\bar{\lambda}^*_{\mathcal{T}}$ is some convex combination of $\hat{\lambda}^*_{\mathcal{T}}$ and $\hat{\lambda}_{\mathcal{T}}$. Thus,

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

From the proof of Proposition B2 and Proposition B3, and recalling that $\mathcal{T}/T \to \pi$, for some $\pi \in (0,1)$,

$$\sqrt{\mathcal{T}} \left(\hat{\phi}^*_{\mathcal{T}} - \phi_0 \right) = \sqrt{\mathcal{T}} \left(\hat{\phi}^*_{\mathcal{T}} - \hat{\phi}_{\mathcal{T}} \right) + \sqrt{\mathcal{T}} \left(\hat{\phi}_{\mathcal{T}} - \phi_0 \right) = O^*_p(1)$$

$$\sqrt{\mathcal{T}} \left(\hat{\theta}^*_{\mathcal{T}} - \theta_0 \right) = \sqrt{\mathcal{T}} \left(\hat{\theta}^*_{\mathcal{T}} - \hat{\theta}_{\mathcal{T}} \right) + \sqrt{\mathcal{T}} \left(\hat{\theta}_{\mathcal{T}} - \theta_0 \right) = O^*_p(1).$$

Hence, by the uniform law of large numbers,

$$\hat{\lambda}^*_{\mathcal{T}} - \lambda_0 = \left(\hat{\lambda}^*_{\mathcal{T}} - \hat{\lambda}_{\mathcal{T}} \right) + \left(\hat{\lambda}_{\mathcal{T}} - \lambda_0 \right) = o^*_p(1) + o_p(1) = o^*_p(1)$$

and, hence,

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

By a similar argument as that in the proof of Proposition B3,

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

Thus, by the same argument used in the proof of Proposition B2 and Proposition B3, we can show that

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

and

$$
\begin{aligned}
&\mathrm{var}^{*}\left(\sqrt{\mathcal{T}}\left(\hat{\lambda}^{*}_{\mathcal{T}}-\hat{\lambda}_{\mathcal{T}}\right)\right)\\
&=\mathrm{var}^{*}\Bigg(\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\psi}^{*}_{\mathcal{T},h}\left(\phi_{0},\theta_{0},\lambda_{0}\right)-\hat{\psi}_{\mathcal{T},h}\left(\phi_{0},\theta_{0},\lambda_{0}\right)\right)\right)\\
&\qquad+\sqrt{\mathcal{T}}F^{\intercal}_{\phi_{0}}\left(\hat{\phi}^{*}_{T}-\hat{\phi}_{T}\right)+\sqrt{\mathcal{T}}F^{\intercal}_{\theta_{0}}\left(\hat{\theta}^{*}_{T}-\hat{\theta}_{T}\right)-\left(\tilde{\boldsymbol{\psi}}^{*}_{\mathcal{T}}-\tilde{\boldsymbol{\psi}}_{\mathcal{T}}\right)\Bigg)\\
&=\mathrm{Avar}\Bigg(\sqrt{\mathcal{T}}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\psi}_{\mathcal{T},h}\left(\phi_{0},\theta_{0},\lambda_{0}\right)-\psi_{0}\left(\phi_{0},\theta_{0},\lambda_{0}\right)\right)\right)\\
&\qquad+\sqrt{\mathcal{T}}F^{\intercal}_{\phi_{0}}\left(\hat{\phi}_{T}-\phi_{0}\right)+\sqrt{\mathcal{T}}F^{\intercal}_{\theta_{0}}\left(\hat{\theta}_{T}-\theta_{0}\right)-\left(\tilde{\boldsymbol{\psi}}_{\mathcal{T}}-\boldsymbol{\psi}\right)\Bigg)+o_{p}(1).
\end{aligned}
$$

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

References

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

Altissimo, F. and A. Mele, 2008, "Simulated Nonparametric Estimation of Dynamic Models," *Review of Economic Studies*, in press.

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, 2005, "What Does the Yield Curve Tell Us about GDP Growth?" *Journal of Econometrics*, forthcoming.

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," forthcoming in *Management Science.*

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.

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. and H. Zhou, 2005, Volatility Puzzles: A Simple Framework for Gauging Return-Volatility Regressions, *Journal of Econometrics,* forthcoming.

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., 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.

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., 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), 803-887.

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

Carr, P. and L. Wu, 2007, "Variance Risk Premia," forthcoming in the *Review of Financial Studies.*

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.

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.

Cox, J.C., J.E. Ingersoll and S.A. Ross, 1985, "A Theory of the Term Structure of Interest Rates," *Econometrica* 53, 385-407.

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

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.

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

Fornari, F. and A. Mele, 2008, "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.

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.

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.

Hull, J. and A. White, 1987, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance* 42, 281-300.

Jiang, G. and Y. Tian, 2005, "The Model-Free Implied Volatility and Its Information Content," *Review of Financial Studies* 18, 1305-1342.

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), 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.

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

Tauchen, G. and V. Todorov, 2008, "Volatility Jumps," working paper Duke University and Northwestern University.

Todorov, V., 2007, "Variance Risk Premium Dynamics," working paper 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 $\phi^\intercal = (\kappa_j, \mu_j, \alpha_j, \beta_j, \bar{\kappa}_j, j = 1, 2)$. 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 average 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 section 3.4.1. 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.

	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 $\theta^{\intercal} = (\kappa_3, \mu_3, \alpha_3, \beta_3, s_0, s_j, j = 1, 2, 3)$, 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 average and standard deviation of stock returns and volatility. The block-bootstrap procedure for the standard errors of the estimates is developed in Section 3.4.2. The sample covers monthly data for the period from January 1950 to December 2006.

	Estimate	Std error
s_0	0.1279	$1.2657 \cdot 10^{-2}$
s_1	0.0998	$3.1345 \cdot 10^{-2}$
s_2	2.5103	$6.1652 \cdot 10^{-1}$
s_3	0.0109	$2.1641 \cdot 10^{-3}$
κ_3	0.0092	$4.6596 \cdot 10^{-3}$
μ_3	1	restr.
α_3	$9.4543 \cdot 10^{2}$	$3.1855 \cdot 10^{2}$
β_3	4.1653	$8.1875 \cdot 10^{-1}$

Table 3

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

$$\begin{aligned}
\boldsymbol{\pi}_1\left(y_1\left(t\right)\right) &= \alpha_1\lambda_{1(1)} + \left(\beta_1\lambda_{1(1)} + \lambda_{2(1)}\right) y_1\left(t\right) \quad \text{(inflation)} \\
\boldsymbol{\pi}_2\left(y_2\left(t\right)\right) &= \alpha_2\lambda_{1(2)} + \left(\beta_2\lambda_{1(2)} + \lambda_{2(2)}\right) y_2\left(t\right) \quad \text{(industrial production)} \\
\boldsymbol{\pi}_3\left(y_3\left(t\right)\right) &= \alpha_3\lambda_{1(3)} + \left(\beta_3\lambda_{1(3)} + \lambda_{2(3)}\right) y_3\left(t\right) \quad \text{(unobservablef factor)}
\end{aligned}$$

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 $\lambda^{\mathsf{T}} = \left(\lambda_{1(1)}, \lambda_{2(1)}, \lambda_{1(2)}, \lambda_{2(2)}, \lambda_{1(3)}, \lambda_{2(3)}\right)$. 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. (36), and (ii) the sample average and standard deviation of the VIX index. The block-bootstrap procedure for the standard errors of the estimates is developed in Section 3.4.3. The sample covers monthly data for the period from January 1990 to December 2006.

	Estimate	Std error
Inflation		
$\lambda_{1(1)}$	6.4605	2.2012
$\lambda_{2(1)}$	-0.1159	0.0109
Ind. Prod.		
$\lambda_{1(2)}$	53.6863	6.4605
$\lambda_{2(2)}$	2.6201	0.6845
Unobs.		
$\lambda_{1(3)}$	2.2079	0.9462
$\lambda_{2(3)}$	3.7559	1.1159

Figures

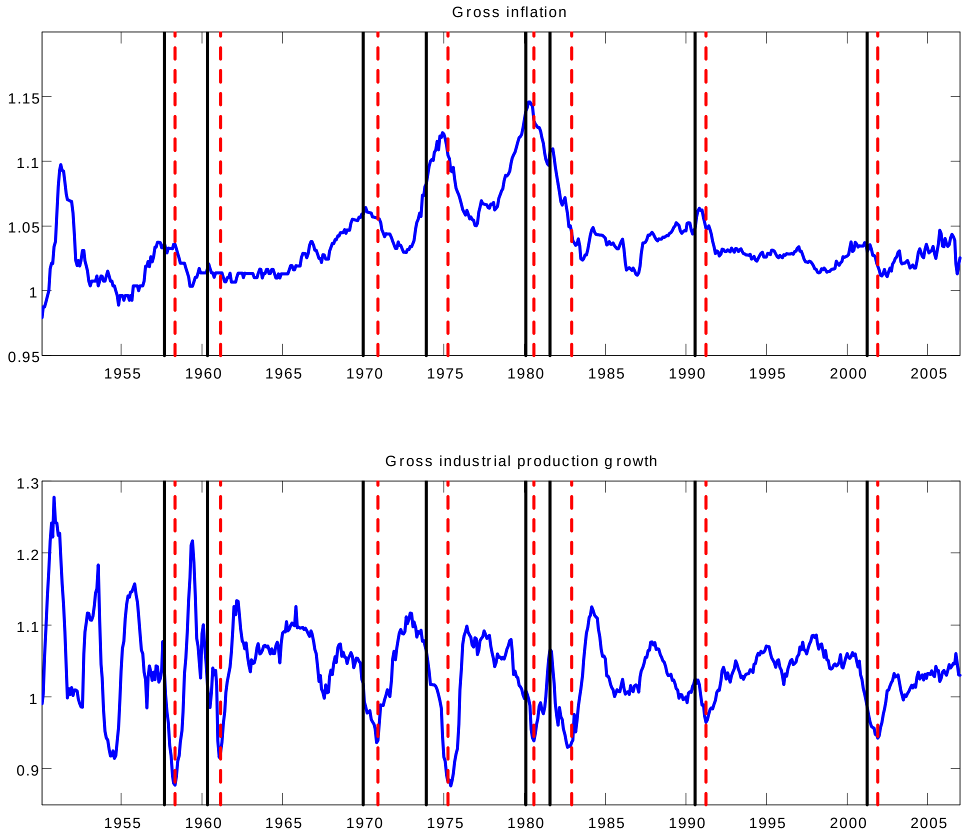

Figure 1 – Industrial production growth and inflation, with NBER dated recession periods. This figure plots the one-year, monthly gross inflation, defined as $y_{1,t} \equiv \text{CPI}_t / \text{CPI}_{t-12}$, and the one-year, monthly gross industrial production growth, defined as $y_{2,t} \equiv \text{IP}_t / \text{IP}_{t-12}$, 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.

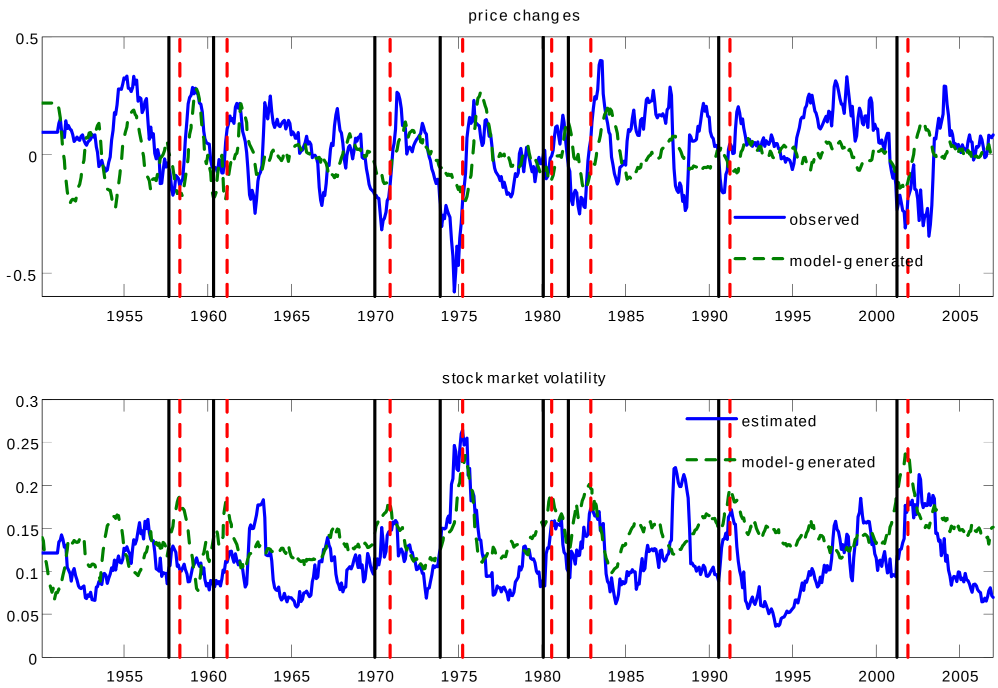

Figure 2 – Returns and volatility along with the model predictions, with NBER dated recession periods. This figure plots one-year ex-post price changes and one-year return volatility, along with their counterparts predicted by the model. The top panel depicts continuously compounded price changes, defined as $R_t \equiv \log(s_t/s_{t-12})$, where s_t is the real stock price as of month t. The middle panel depicts smoothed return volatility, defined as $\text{Vol}_t \equiv \sqrt{6\pi} \cdot 12^{-1} \sum_{i=1}^{12} |\log(s_{t+1-i}/s_{t-i})|$, 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 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.

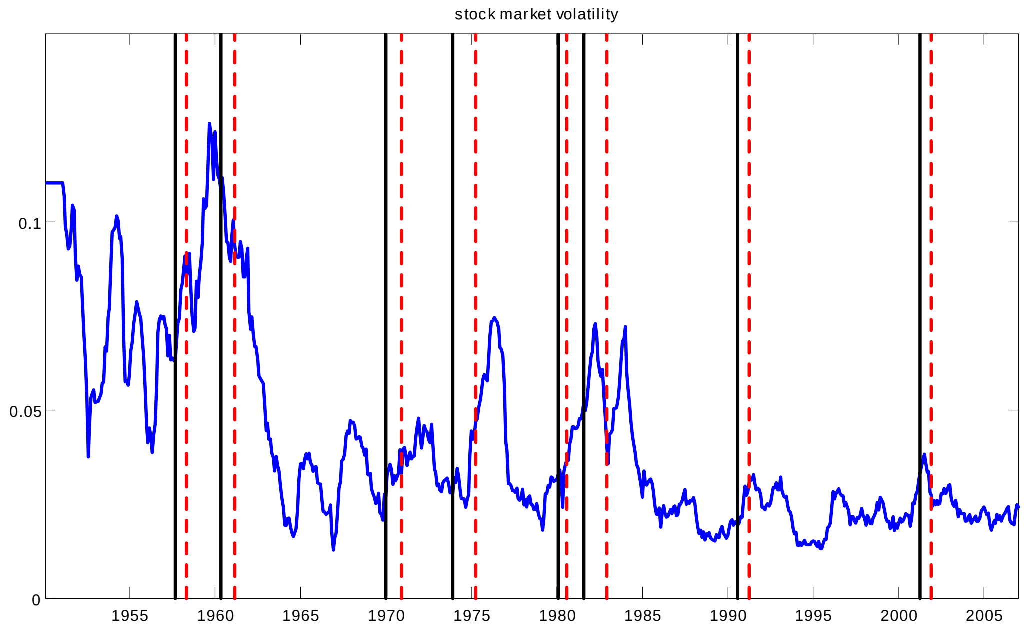

Figure 3 – Return volatility along with model predictions obtained without the unobservable factor, with NBER dated recession periods. This figure plots return the volatility predicted by the model, obtained as $\text{Vol}_t \equiv \sqrt{6\pi} \cdot 12^{-1} \sum_{i=1}^{12} |\log(s_{t+1-i}/s_{t-i})|$, where s_t is the real stock price as of month t. 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 feeezing 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.

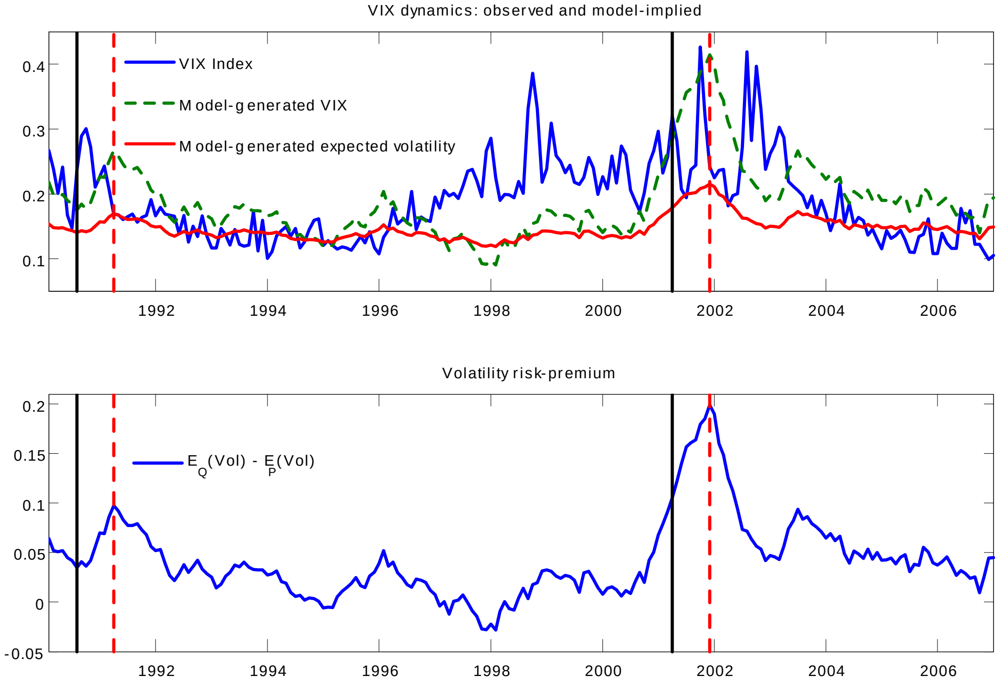

Figure 4 – The VIX Index and volatility risk-premia, with NBER dated recession periods. This figure plots the VIX index along with model's predictions. The top panel depicts (i) the VIX index, (ii) the VIX index predicted by the model, and (iii) the VIX index predicted by the model in an economy without 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, $\sigma^2\left(\boldsymbol{y}\right)$ 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 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.

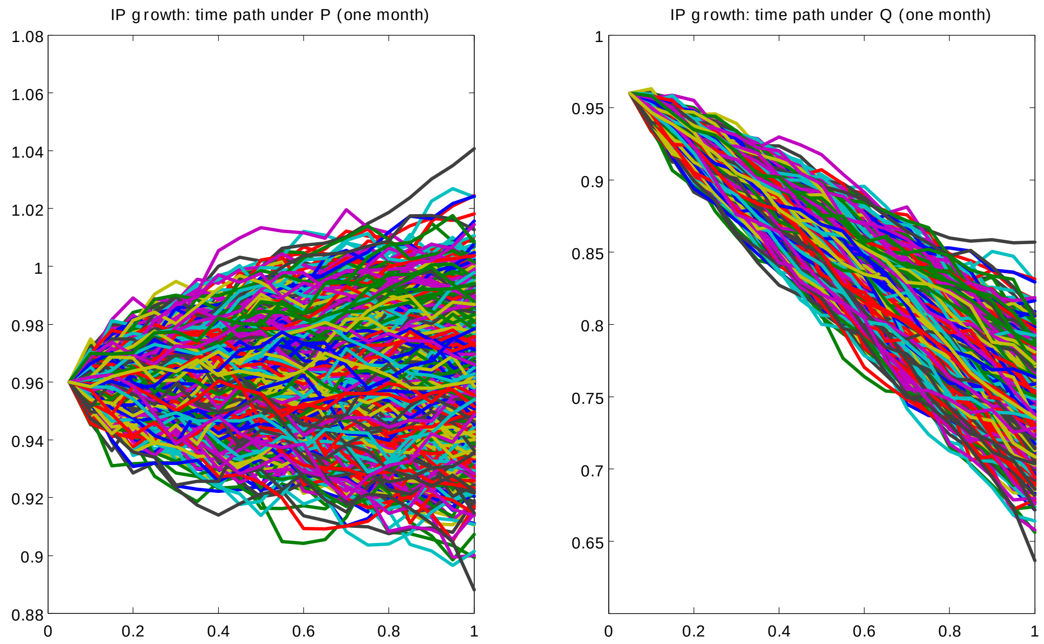

Figure 5 – Sample paths of industrial production growth. This figure plots 1000 simulations of one month paths of the gross industral production growth, with starting values fixed at 1.03 (gross inflation) and 0.96 (gross industrial production growth). The left panel displays the sample paths under P, the physical probability. The right panel depicts the sample paths under Q, the risk-neutral probability. Simulations are performed with parameters set equal to the point estimates reported in Tables 1 and 2.

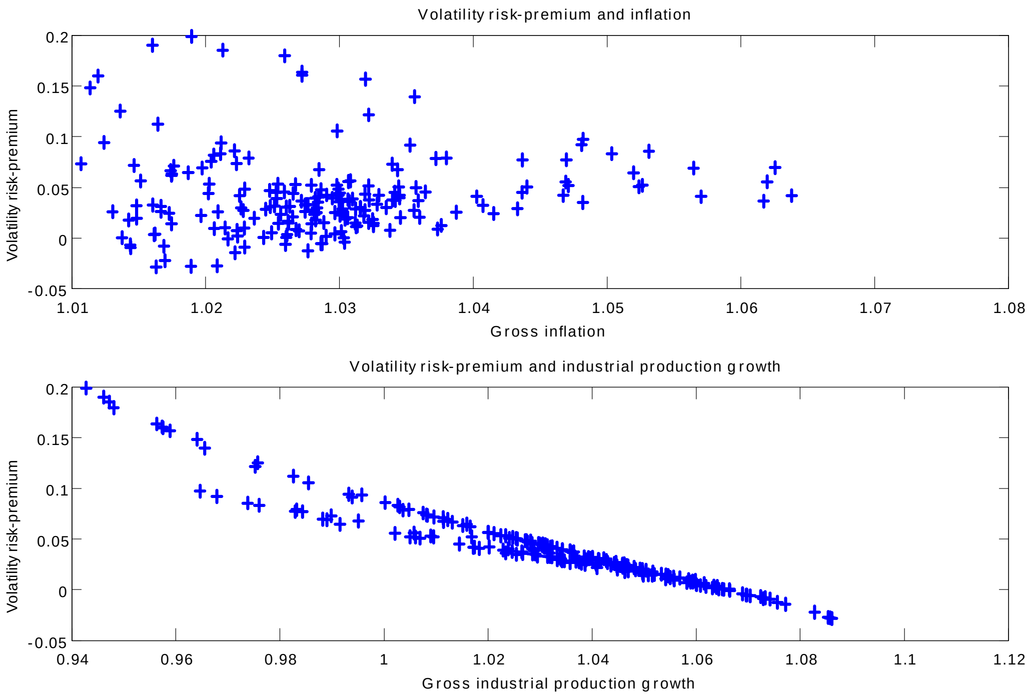

Figure 6 – Volatility risk-premium against inflation and industrial production growth. This figure provides scatterplots of the volatility risk-premium predicted by the model, depicted in Figure 3 (bottom panel), against the two macroeconomic factors depicted in Figure 1 (inflation and growth). Each 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.