Macroeconomic Determinants of Stock Market Returns, Volatility and Volatility Risk-Premia

Valentina Corradi
Walter Distaso
Antonio Mele

DISCUSSION PAPER NO 616

DISCUSSION PAPER SERIES

June 2008

Valentina Corradi is Professor of Econometrics at the University of Warwick. Her current research interests include (i) modelling, estimation and forecasting of volatility models via realized measures, (ii) predictive density evaluation, (iii) bootstrap refinements, (iv) simulation-based inference. She is Associate Editor of the Review of Economic Studies and of the Econometrics Journal, and Fellow of the Journal of Econometrics. Her work has been published by leading journals in economics and econometrics such as International Economic Review, Review of Economic Studies, Journal of Econometrics, Journal of Economic Theory, Journal of Business Economic and Statistics, and Econometric Theory. Walter Distaso is at Tanaka Business School from 2006. He holds a PhD from the University of York. Previously, he held positions at the University of Exeter and Queen Mary. He has also been a visiting professor at the IMF. His research interests are in the area of (i) estimation, specification testing and prediction of financial Volatility in continuous time models (ii) analyzing time series using long memory models (iii) identifying the macroeconomic determinants of stock-market volatility (iv) studying the dependence of multivariate financial time series using copulas (v) evaluating competing trading strategies (vi) analyzing the features and the effects of market microstructure noise. Antonio Mele is a Reader in Finance at the London School of Economics and an Academic Research Fellow at the Financial Markets Group. His academic interests include empirical and theoretical research into (i) the nexus between capital markets and the macro-economy (ii) derivative pricing and (iii) financial econometrics. His research is published by leading journals in finance and economics such as the Review of Financial Studies, the Journal of Financial Economics and the Review of Economic Studies, among others.

Macroeconomic Determinants of Stock Market Returns, Volatility and Volatility Risk-Premia[*]

Valentina Corradi
University of Warwick

Walter Distaso
Tanaka Business School

Antonio Mele
London School of Economics

First draft: July 22, 2005
This draft: June 10, 2008
Preliminary draft

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 return 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: realized volatility; volatility risk-premium; macroeconomic factors; no arbitrage restrictions; concentrated simulated general method of moments, block-bootstrap.

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 variance risk-premium (see, for example, Carr and Wu (2004) and Bakshi and Madan (2006)). The variance risk-premium is defined as the difference between the expectation of future stock market volatility under the true and the risk-neutral probability. It quantifies how much a representative agent is willing to pay to ensure that volatility will not raise above a given threshold. Thus, it is a very intuitive and general measure of risk-aversion. Previous important work by Bollerslev and Zhou (2005) and Bollerslev, Gibson and Zhou (2004) has analyzed how this variance risk-premium is related to a number of macroeconomic factors. The authors regressed semi-parametric measures of the variance risk-premium on these 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 macroeconomic factors and risk-premia is richer than in the previous

papers as it explicitly accounts for the necessary no-arbitrage relations that link asset prices and, hence, return volatility, to macroeconomic factors.

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 price 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, but do not directly related these to markets, preferences and technology.

Our model works quite simply. We start with exogenously specifying the joint dynamics of both macroeconomic and latent factors. Then, we assume that dividends and risk-premia are essentially affine functions of the factors, along the lines of Duffee (2002). We show that the resulting no-arbitrage stock price is affine in the factors. Our model is also related to previous approaches in the literature. For example, Bekaert and Grenadier (2001) and Ang and Liu (2004) formulated discrete-time models in which the key pricing factors are exogenously given. Furthermore, Mamaysky (2002) derived a continuous-time model based on an exogenous specification of the price-dividend ratio. There are important differences between these models and ours. First, our model is in continuous-time and thus avoids theoretical inconsistencies arising in the discrete time setting considered by Bekaert and Grenadier (2001). Second, a continuous-time setting is particularly appealing given our objective to estimate volatility and volatility risk-premia through measures of realized volatility. Third, Ang and Liu (2004) consider a discrete-time setting in which expected returns are exogenous to their model; in our model, expected returns are endogenous. Finally, our model works differently from Mamaysky's because it endogenously determines the price-dividend ratio.

Estimating our model is challenging. In our model, volatility is endogenous, which makes parameters' identification a quite delicate issue. The main difficulty we face is that return volatility arises by a rational price formation process. 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 identifiability 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 published 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 exploit the general ideas underlying the "realized volatility" literature to implement consistent estimators of the VIX index (see, e.g., Barndorff-Nielsen and Shephard (2007) for a survey on realized volatility). 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 our no-arbitrage model for the stock price, return volatility and the variance risk-premia. Section 3 illustrates the estimation strategy. Section 4 presents our empirical results, 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,

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

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,

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

4

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

While we do not necessarily observe every single component of $\boldsymbol{y}\left(t\right)$, we do observe discretely sampled paths of macroeconomic variables such as industrial production, unemployment or inflation. Let $\{\mathrm{M}_j\left(t\right)\}_{t=1,2,\cdots}$ be the discretely sampled path of the macroeconomic variable $\mathrm{M}_j\left(t\right)$ where, for example, $\mathrm{M}_j\left(t\right)$ 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:

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

where the collection of functions $\{\varphi_j\}_j$ determines how the factors dynamics impinge upon the evolution of the overall macroeconomic conditions. 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 $\boldsymbol{y}\left(t\right)$ 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. In other words, and 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).

Formally, we assume that there exists a rational pricing function $s\left(\boldsymbol{y}\left(t\right)\right)$ such that the real stock price at time t, $s\left(t\right)$ say, is $s\left(t\right) \equiv s\left(\boldsymbol{y}\left(t\right)\right)$. We let this price function be twice continuously differentiable in $\boldsymbol{y}$. (Given the assumptions and conditions we give below, this differentiability condition holds in our model.) By Itô's lemma, $s\left(t\right)$ satisfies,

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

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

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

Next, we model the pricing kernel in the economy. The Radon-Nikodym derivative of Q, the equivalent martingale measure, with respect to P on $\mathbb{F}(T)$ is,

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

for some adapted $\boldsymbol{\Lambda}(t)$, the risk-premium process. 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. 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 lead to the same affine pricing function in Proposition 1 below, but statistical inference for the resulting model would be hindered. Moreover, interest rate volatility appears to play a limited role in the main applications we consider in this paper.

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

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

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

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, 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,[1] and assuming no-bubbles, Eq. (5) implies that the stock price is,

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

where $\mathbb{E}$ is the expectation taken under the equivalent martingale measure. 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 $\delta\left(\cdot\right)$ is affine in $\boldsymbol{y}$ implies that the stock price is also affine in $\boldsymbol{y}$. Precisely, let

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

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, (i) eq. (7) holds, and (ii) the rational stock function $s\left(\boldsymbol{y}\right)$ is linear in the state vector $\boldsymbol{y}$, viz*

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

where

$$\boldsymbol{c} = \boldsymbol{\kappa}\boldsymbol{\mu} - \boldsymbol{\Sigma}\left(\begin{array}{ccc} \alpha_1\lambda_{1(1)} & \cdots & \alpha_d\lambda_{1(d)} \end{array}\right)^\top,$$

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

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

Proposition 1 allows us to describe what this model predicts in terms of no-arbitrage restrictions between stochastic volatility and risk-premia. In particular, use Eq. (9) to compute volatility through Eq. (4). We obtain,

$$\sigma^2\left(t\right) = \frac{\sqrt{\left\|\boldsymbol{\delta}^\top\left(\boldsymbol{D} + r\boldsymbol{I}_{n\times n}\right)^{-1}\boldsymbol{\Sigma}\boldsymbol{V}\left(\boldsymbol{y}\left(t\right)\right)\right\|^2}}{\frac{\delta_0 + \boldsymbol{\delta}^\top\left(\boldsymbol{D} + r\boldsymbol{I}_{n\times n}\right)^{-1}\boldsymbol{c}}{r} + \boldsymbol{\delta}^\top\left(\boldsymbol{D} + r\boldsymbol{I}_{n\times n}\right)^{-1}\boldsymbol{y}\left(t\right)}. \tag{10}$$

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,

---

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

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,

$$\begin{cases} \dfrac{\mathrm{d}s\left(t\right)}{s\left(t\right)} = m\left(t\right)\mathrm{d}t + v\left(t\right)\mathrm{d}W_1\left(t\right) \\ \mathrm{d}v\left(t\right)^2 = \kappa\left(\mu - v\left(t\right)^2\right)\mathrm{d}t + \sigma v\left(t\right)\left(\rho\mathrm{d}W_1\left(t\right) + \sqrt{1-\rho^2}\mathrm{d}W_2\left(t\right)\right) \end{cases} \qquad (11)$$

for some adapted process $m\left(t\right)$ 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 very important, 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. (10) 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.[2] The new index reflects recent advances into the option pricing literature. Given an asset price process $s\left(t\right)$ that is continuous in time (as for the asset price of our model in Eq. (9)), and all available information $\mathbb{F}\left(t\right)$ at time t, define the integrated return variance on a given interval $[t, T]$ as,

$$IV_{t,T} = \int_t^T \left(\left.\frac{\mathrm{d}}{\mathrm{d}\tau} var\left[\log s\left(\tau\right)\right| \mathbb{F}\left(u\right)]\right|_{\tau=u}\right)\mathrm{d}u. \qquad (12)$$

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

---

[2]If the interest rate is zero, then, in the absence of arbitrage opportunities, the variance swap rate is simply the expectation of the future integrated return volatility under the risk-neutral probability, as defined in Eq. (12) below.

of the future integrated variance is a functional of put and call options written on the asset:

$$\mathbb{E}\left[IV_{t,T}|\,\mathbb{F}\left(t\right)\right] = 2\left[\int_0^{F(t)} \frac{P\left(t,T,K\right)}{u\left(t,T\right)}\frac{1}{K^2}\mathrm{d}K + \int_{F(t)}^{\infty} \frac{C\left(t,T,K\right)}{u\left(t,T\right)}\frac{1}{K^2}\mathrm{d}K\right], \qquad (13)$$

where $F\left(t\right) = u\left(t,T\right)s\left(t\right)$ is the forward price, $C\left(t,T,K\right)$ and $P\left(t,T,K\right)$ are the prices as of time t of a call and a put option expiring at T and struck at K, and $u\left(t,T\right)$ is the price as of time t of a pure discount bond expiring at T. A variance swap is a contract with payoff proportional to the difference between the realized integrated variance, (12), and some strike price, the variance swap rate. In the absence of arbitrage opportunities, the variance swap rate is then given by Eq. (13).

Eq. (13) is helpful because it relies on a nonparametric method to compute the risk-neutral expectation of the integrated variance. Our model predicts that the risk-neutral expectation of the integrated variance is:

$$\mathbb{E}\left[IV_{t,T}|\,\mathbb{F}\left(t\right)\right] = \int_t^T \mathbb{E}[\sigma\left(u\right)^2 \mid \mathbb{F}\left(t\right)]\mathrm{d}u,$$

where $\mathbb{F}\left(t\right)$ is the filtration generated by the multidimensional Brownian motion in Eq. (1), and $\sigma\left(t\right)^2$ is given in Eq. (10). 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. (13).

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

$$\mathrm{VRP}\left(t\right) \equiv \sqrt{\frac{1}{T-t}}\left(\sqrt{\mathbb{E}\left[IV_{t,T}|\,\mathbb{F}\left(t\right)\right]} - \sqrt{E\left[IV_{t,T}|\,\mathbb{F}\left(t\right)\right]}\right), \qquad (14)$$

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

2.5   The leading model

We formulate a few specific assumptions to make the model amenable to empirical work. First, we assume that two macroeconomic aggregates, inflation and industrial production growth, are the only observable factors (say y_1 and y_2) affecting 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 $\mathrm{M}_1\left(t\right)$ is the consumer price index as of month t and $\mathrm{M}_2\left(t\right)$ 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 $\varphi_j\left(\boldsymbol{y}\right) \equiv \log y_j$.

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),

$$
\boldsymbol{\kappa} = \left[\begin{array}{ccc} \kappa_1 & \bar{\kappa}_1 & 0 \\ \bar{\kappa}_2 & \kappa_2 & 0 \\ 0 & 0 & \kappa_3 \end{array} \right],
$$

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, in the absence of feedbacks, and $\bar{\kappa}_1$ and $\bar{\kappa}_2$ are the feedback parameters. Moreover, we take $\boldsymbol{\Sigma} = \boldsymbol{I}_{3\times3}$ and the vectors $\boldsymbol{\beta}_i$ so as to make y_j solution to,

$$
\mathrm{d}y_j(t) = \left[\kappa_j \left(\mu_j - y_j(t) \right) + \bar{\kappa}_j \left(\bar{\mu}_j - \bar{y}_j(t) \right) \right] \mathrm{d}t + \sqrt{\alpha_j + \beta_j y_j(t)} \mathrm{d}W_j(t), \quad j = 1, 2, 3, \quad (15)
$$

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,

$$
\boldsymbol{\pi}(\boldsymbol{y}) \equiv \boldsymbol{\Sigma} \boldsymbol{V}(\boldsymbol{y}) \boldsymbol{\Lambda}(\boldsymbol{y}) = \left(\begin{array}{c} \alpha_1 \lambda_{1(1)} + \left(\beta_1 \lambda_{1(1)} + \lambda_{2(1)} \right) y_1 \\ \alpha_2 \lambda_{1(2)} + \left(\beta_2 \lambda_{1(2)} + \lambda_{2(2)} \right) y_2 \\ \alpha_3 \lambda_{1(3)} + \left(\beta_3 \lambda_{1(3)} + \lambda_{2(3)} \right) y_3 \end{array} \right) \quad (16)
$$

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 (8) satisfies,

$$
\delta(\boldsymbol{y}) = \delta_0 + \delta_1 y_1 + \delta_2 y_2 + \delta_3 y_3. \quad (17)
$$

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

$$
s(\boldsymbol{y}) = s_0 + \sum_{j=1}^{3} s_j y_j, \quad (18)
$$

where

$$s_0 = \frac{1}{r} \left[\delta_0 + \sum_{j=1}^{3} s_j \left(\kappa_j \mu_j + \bar{\kappa}_j \bar{\mu}_j - \alpha_j \lambda_{1(j)} \right) \right] \tag{19}$$

$$s_j = \frac{\delta_j \left(r + \kappa_i + \lambda_{1(i)} \beta_i + \lambda_{2(i)} \right) - \delta_i \bar{\kappa}_i}{\prod_{h=1}^{2} \left(r + \kappa_h + \lambda_{1(h)} \beta_h + \lambda_{2(h)} \right) - \bar{\kappa}_1 \bar{\kappa}_2}, \quad \text{for } j, i \in \{1, 2\} \text{ and } i \neq j \tag{20}$$

$$s_3 = \frac{\delta_3}{r + \kappa_3 + \lambda_{1(3)} \beta_3 + \lambda_{2(3)}} \tag{21}$$

and where $\bar{\kappa}_j$ and $\bar{\mu}_j$ are as in Eq. (15).

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. (16) is to changes in the state process y. Instead, the parameters δ_i determine how sensitive the dividend process in Eq. (17) is to changes in 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

Our estimation strategy relies on a three-step procedure. In the first step, we estimate the parameters of the two-dimensional affine diffusion describing the macroeconomic factors dynamics, that is we estimate $\phi^{\mathsf{T}} = \left(\kappa_j, \mu_j, \alpha_j, \beta_j, \bar{\kappa}_j, j = 1, 2 \right)$. In the second step, we estimate the reduced form parameters linking the equilibrium stock price to the factors, as in Eq. (18), as well as the parameters of a restricted version of the affine diffusion describing the latent factor dynamics, that is we estimate $\theta^{\mathsf{T}} = \left(\kappa_3, \mu_3, \alpha_3, \beta_3, s_0, s_j, j = 1, 2, 3 \right)$ imposing $\mu_3 = 1$. In the third step, we estimate the risk premia parameters $\lambda^{\mathsf{T}} = \left(\lambda_{1(1)}, \lambda_{1(2)}, \lambda_{1(3)}, \lambda_{2(1)}, \lambda_{2(2)}, \lambda_{2(3)} \right)$ using a functional approximation of the model implied VIX, based on the parameters estimated in the previous two steps, and the model-free VIX series.

As at any step we do not have a closed form expression of either the likelihood function or sets of moment conditions, we rely on a simulation based approach. Broadly speaking, our strategy can be viewed as an hybrid of Indirect Inference (Gourieroux, Monfort and Renault, 1993) and Simulated Generalized Method of Moments (Duffie and Singleton, 1993). In fact, we aim at matching impulse response functions as well as sample moments of historical and simulated data.

3.1 Moment conditions and parameter estimation for the macroeconomic factors

Using a Milhstein approximation scheme of the diffusion in Eq. (15), with a discrete interval Δ, we simulate H paths of length T for the two observable factors, and sample them at the same frequency of the data, say $y_{1,t,h}^{\Delta}(\phi), y_{2,t,h}^{\Delta}(\phi)$, $h = 1, \cdots, H$, $t = 1, \cdots, T$. We then estimate a VAR model based on historical and simulated data, i.e. for $i = 1, 2$ we estimate the following auxiliary models,

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

$$y_{i,t,h}^{\Delta}(\phi) = \varphi_{i,1,h} + \sum_{j \in \{12,24\}} \varphi_{i,1,j,h} y_{1,t-j,h}^{\Delta}(\phi) + \sum_{j \in \{12,24\}} \varphi_{i,2,j,h}(\phi) y_{2,t-j,h}^{\Delta}(\phi) + \epsilon_{y_h,i,t}(\phi)$$

Let $\mathbf{Y}_t = (y_{1,t-12}, y_{1,t-24}, y_{2,t-12}, y_{2,t-24})^{\mathsf{T}}$, and let $\mathbf{Y}_{t,h}^{\Delta}(\phi)$ be the simulated counterpart, obtained replacing $y_{1,t}$ with $y_{1,t,h}^{\Delta}(\phi)$, for each parameter trial ϕ. Define the OLS estimators constructed using actual and simulated data as,

$$\tilde{\varphi}_{i,T} = \left(\frac{1}{T} \sum_{t=25}^{T} \mathbf{Y}_t \mathbf{Y}_t^{\mathsf{T}}\right)^{-1} \frac{1}{T} \sum_{t=25}^{T} \mathbf{Y}_t y_{i,t} \tag{22}$$

$$\hat{\varphi}_{i,T,h}^{\Delta}(\phi) = \left(\frac{1}{T} \sum_{t=25}^{T} \mathbf{Y}_{t,h}^{\Delta}(\phi) \mathbf{Y}_{t,h}^{\Delta}(\phi)^{\mathsf{T}}\right)^{-1} \frac{1}{T} \sum_{t=25}^{T} \mathbf{Y}_{t,h}^{\Delta}(\phi) y_{i,t,h}^{\Delta}(\phi).$$

Finally, let

$$\tilde{\varphi}_T = \left(\tilde{\varphi}_{1,T}, \tilde{\varphi}_{2,T}, \bar{y}_1, \bar{y}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2\right)^{\mathsf{T}} \tag{23}$$

$$\hat{\varphi}_{i,T,h}^{\Delta}(\phi) = \left(\hat{\varphi}_{1,T,h}^{\Delta}(\phi), \hat{\varphi}_{2,T,h}^{\Delta}(\phi), \bar{y}_{1,h}^{\Delta}(\phi), \bar{y}_{1,h}^{\Delta}(\phi), \hat{\sigma}_{1,h}^{2,\Delta}(\phi), \hat{\sigma}_{2,h}^{2,\Delta}(\phi)\right)$$

where for $i = 1, 2$ $\bar{y}_i, \hat{\sigma}_i^2, \bar{y}_{i,h}^{\Delta}(\phi), \hat{\sigma}_{i,h}^{2,\Delta}(\phi)$ are sample mean and variance based on actual and simulated data. The estimator of the macroeconomic factor parameters and its probability limit are then given by:

$$\hat{\phi}_T = \arg\min_{\varphi} \left(\left(\frac{1}{H} \sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi) - \tilde{\varphi}_T\right)^{\mathsf{T}} \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi) - \tilde{\varphi}_T\right)\right),$$

$$\phi_0 = \arg\min_{\varphi} \text{p} \lim_{T \to \infty, \Delta \to 0} \left(\left(\frac{1}{H} \sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi) - \tilde{\varphi}_T\right)^{\mathsf{T}} \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\varphi}_{T,h}^{\Delta}(\phi) - \tilde{\varphi}_T\right)\right).$$

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

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

where

$$\mathbf{V}_1 = \left(1 + \frac{1}{H}\right)(\mathbf{D}_1^\intercal \mathbf{D}_1)^{-1}\mathbf{D}_1^\intercal \mathbf{J}_1 \mathbf{D}_1 (\mathbf{D}_1^\intercal \mathbf{D}_1)^{-1}$$

$$\mathbf{D}_1 = \mathrm{p}\lim \nabla_\phi \left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}^\Delta (\phi_0)\right)$$

$$\mathbf{J}_1 = \mathrm{Avar}\left(\sqrt{T}\left(\tilde{\varphi}_T - \varphi_0\right)\right) = \mathrm{Avar}\left(\sqrt{T}\left(\hat{\varphi}_{T,h}^\Delta (\phi_0) - \varphi_0\right)\right), \text{ for all } h.$$

3.2 Moment conditions and parameter estimation for the asset price and the unobserved factor

Proposition 1 establishes that the equilibrium stock price is an affine function of the factors, i.e.

$$s_t = s(y_t) = s_0 + \sum_{j=1}^{3} s_j y_j(t),$$

where the parameters $s_0, s_j,\ j = 1, 2, 3$ are functions of the structural parameters, as established in Eqs. (19)-(21). Using data on macroeconomic factors and stock returns, we cannot identify all the structural parameters. In particular, we cannot identify both dividends and risk premia parameters, as there are infinite combinations of δ and λ giving raise to the same equilibrium stock price.

As we are not interested in the dividends parameters per se, we proceed by estimating the reduced form parameter (s_0, s_1, s_2, s_3) as well as the parameters of the latent factor $y_{3,t}$, i.e. $(\kappa_3, \mu_3, \alpha_3, \beta_3)$. However, as the latent factor is independent of the observable factors, samples on stock returns and macroeconomic factors do not allow us to separately identify s_3 and $(\kappa_3, \mu_3, \alpha_3, \beta_3)$. We impose the restriction $\mu_3 = 1$, and define a new factor $Z(t) = s_3 y_3(t)$, as

$$\mathrm{d}Z(t) = \kappa_3\left(A - Z(t)\right)\mathrm{d}t + \sqrt{B + CZ(t)}\mathrm{d}W_{3,t},$$

where $A = \mu_3 s_3 = s_3$, $B = \alpha_3 s_3^2$, $C = \beta_3 s_3$. We simulate H paths of length T for the (new) unobservable factor, using a discrete interval Δ, and sample it at the same frequency of the data, i.e. $Z_{t,h}(\theta_u)$, $h = 1, \cdots, H$, $t = 1, \cdots, T$, and $\theta_u = (\kappa_3, \alpha_3, \beta_3, s_3)$. We then construct the simulated stock price process as

$$s_{t,h}^\Delta (\theta) = s_0 + s_1 y_{1,t} + s_2 y_{2,t} + Z_{t,h}^\Delta (\theta_u), \tag{24}$$

with $\theta = (\kappa_3, \alpha_3, \beta_3, s_0, s_1, s_2, s_3)$, and $s_0 = \bar{s}_h^\Delta (\theta) - s_1\bar{y}_1 - s_2\bar{y}_2 - \bar{Z}_3^\Delta (\theta_u)$, where $\bar{s}_h^\Delta (\theta)$, $\bar{y}_1$, $\bar{y}_2$ and $\bar{Z}_3^\Delta (\theta_u)$ are the sample means of $s_{t,h}^\Delta (\theta), y_{1,t}, y_{2,t}$ and $Z_{3,t}^\Delta (\theta_u)$. Note that the stock price has been simulated using the observed samples for $y_{1,t}$ and $y_{2,t}$.

Following Mele (2007) and Fornari and Mele (2006), we measure the volatility of the monthly continuously compounded price changes, $R_t = \log\left(s_t/s_{t-1}\right)$, as,

$$\text{Vol}_t = \sqrt{6\pi} \cdot \frac{1}{12} \sum_{i=1}^{12} |R_{t+1-i}|. \tag{25}$$

Hereafter, let $R_{t,h}^\Delta(\theta) = \log\left(s_{t,h}^\Delta(\theta)/s_{t-1,h}^\Delta(\theta)\right)$ and $\text{Vol}_{t,h}^\Delta(\theta) = \sqrt{6\pi} \cdot \frac{1}{12} \sum_{i=1}^{12} \left|R_{t+1-i,h}^\Delta(\theta)\right|$ be simulated counterparts of R_t and Vol_t. In the sequel, we rely on the following two auxiliary models,

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

$$\text{Vol}_t = a^\text{V} + \sum_{i\in\{6,12,18,24,36,48\}} \phi_i \text{Vol}_{t-i} + \sum_{i\in\{12,24,36,48\}} b_{1,i}^\text{V} y_{1,t-i} + \sum_{i\in\{12,24,36,48\}} b_{2,i}^\text{V} y_{2,t-i} + \epsilon_{h,t}^\text{V},$$

and their simulation-based counterparts,

$$R_{t,h}^\Delta(\theta) = a_h^\text{R} + b_{1,12,h}^\text{R} y_{1,t-12} + b_{2,12,h}^\text{R} y_{2,t-12} + \epsilon_t^\text{R},$$

$$\text{Vol}_{t,h}^\Delta(\theta) = a_h^\text{V} + \sum_{i\in\{6,12,18,24,36,48\}} \phi_{i,h} \text{Vol}_{t,h}^\Delta(\theta) + \sum_{i\in\{12,24,36,48\}} b_{1,i,h}^\text{V} y_{1,t-i} + \sum_{i\in\{12,24,36,48\}} b_{2,i,h}^\text{V} y_{2,t-i} + \epsilon_{h,t}^\text{V},$$

Let $\mathbf{Y}_t = (y_{1,t-12}, y_{2,t-12})^\intercal$, $\mathbf{V}_t = (y_{1,t-i}, y_{2,t-i}, \text{Vol}_{t-6}, \text{Vol}_{t-i}, \ i \in \{12,18,24,36,48\})^\intercal$ and $\mathbf{V}_{t,h}^\Delta(\theta) = \left(y_{1,t-i}, y_{2,t-i}, \text{Vol}_{t,h}^\Delta(\theta)_{t-6}, \text{Vol}_{t,h}^\Delta(\theta)_{t-i}, \ i \in \{12,18,24,36,48\}\right)^\intercal$, and define the OLS estimators constructed using actual and simulated data as,

$$\tilde{\vartheta}_{1,T} = \left(\tfrac{1}{T} \sum_{t=13}^T \mathbf{Y}_t \mathbf{Y}_t^\intercal\right)^{-1} \tfrac{1}{T} \sum_{t=13}^T \mathbf{Y}_t R_t$$

$$\tilde{\vartheta}_{2,T} = \left(\tfrac{1}{T} \sum_{t=49}^T \mathbf{V}_t \mathbf{V}_t^\intercal\right)^{-1} \tfrac{1}{T} \sum_{t=49}^T \mathbf{V}_t \text{Vol}_t$$

and

$$\hat{\vartheta}_{1,T,h}^\Delta(\theta) = \left(\tfrac{1}{T} \sum_{t=13}^T \mathbf{Y}_t \mathbf{Y}_t^\intercal\right)^{-1} \tfrac{1}{T} \sum_{t=13}^T \mathbf{Y}_t R_{t,h}^\Delta(\theta)$$

$$\hat{\vartheta}_{2,T,h}^\Delta(\theta) = \left(\tfrac{1}{T} \sum_{t=49}^T \mathbf{V}_{t,h}^\Delta(\theta) \mathbf{V}_{t,h}^\Delta(\theta)^\intercal\right)^{-1} \tfrac{1}{T} \sum_{t=49}^T \mathbf{V}_{t,h}^\Delta(\theta) \text{Vol}_{t,h}^\Delta(\theta) \tag{26}$$

and finally, define,

$$\tilde{\vartheta}_T = \left(\tilde{\vartheta}_{1,T}, \tilde{\vartheta}_{2,T}, \bar{R}, \overline{\text{Vol}}\right)^\intercal \quad ; \quad \hat{\vartheta}_{T,h}^\Delta(\theta) = \left(\hat{\vartheta}_{1,T,h}^\Delta(\theta), \hat{\vartheta}_{2,T,h}^\Delta(\theta), \bar{R}_h^\Delta(\theta), \overline{\text{Vol}}_h^\Delta(\theta)\right)^\intercal,$$

where $\bar{R}$, $\overline{\text{Vol}}$, $\bar{R}_h^\Delta(\theta)$, $\overline{\text{Vol}}_h^\Delta(\theta)$ are sample mean of return and volatility based on actual and simulated data. The estimator of the parameter of interest θ and its probability limit are then given by:

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

14

and

$$\theta_0 = \arg\min_{\theta} \underset{T \to \infty, \Delta \to 0}{\text{p lim}} \left(\left(\frac{1}{H} \sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{\Delta}(\theta) - \tilde{\vartheta}_T \right)^{\mathsf{T}} \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{\Delta}(\theta) - \tilde{\vartheta}_T \right) \right).$$

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

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

where

$$\mathbf{V}_2 = \left(1 + \frac{1}{H} \right) (\mathbf{D}_2^{\mathsf{T}} \mathbf{D}_2)^{-1} \mathbf{D}_2^{\mathsf{T}} (\mathbf{J}_2 - \mathbf{K}_2) \mathbf{D}_2 (\mathbf{D}_2^{\mathsf{T}} \mathbf{D}_2)^{-1}$$

$$\mathbf{D}_2 = \text{p lim} \, \nabla_{\theta} \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{\Delta}(\theta_0) \right)$$

$$\mathbf{J}_2 = \text{Avar} \left(\sqrt{T} \left(\tilde{\vartheta}_T - \vartheta_0 \right) \right) = \text{Avar} \left(\sqrt{T} \left(\hat{\vartheta}_{T,h}^{\Delta}(\theta_0) - \vartheta_0 \right) \right), \text{ for all } h$$

$$\mathbf{K}_2 = \text{Acov} \left(\sqrt{T} \left(\tilde{\vartheta}_T - \vartheta_0 \right), \sqrt{T} \left(\hat{\vartheta}_{T,h}^{\Delta}(\theta_0) - \vartheta_0 \right)^{\mathsf{T}} \right) \text{ for all } h$$

$$= \text{Acov} \left(\sqrt{T} \left(\hat{\vartheta}_{T,h'}^{\Delta}(\theta_0) - \vartheta_0 \right), \sqrt{T} \left(\hat{\vartheta}_{T,h}^{\Delta}(\theta_0) - \vartheta_0 \right)^{\mathsf{T}} \right), \text{ for all } h \neq h'.$$

Note that the structure of the asymptotic covariance matrix is different from that in Lemma 1,; 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. In fact, we simulate the stock price process conditionally on the sample realizations for the observable factors, thus performing conditional (simulated) inference. It is immediate to see that the use of observed values of $y_{1,t}$ and $y_{2,t}$ in (24), provides an efficiency improvement over unconditional (simulated) inference.

3.3 Identification and estimation of the risk-premium parameters

It remains to estimate the risk premia parameters. We now construct a functional approximation of the model-implied VIX index, using the parameters estimated in the previous two steps. Then, we use actual samples of the model-free VIX index, as published by the Chicago Board of Exchange (CBOE), and obtain estimates of λ by matching the impulse response functions as well as other sample moments of the VIX index implied by the model and its model-free counterpart.

Let $W(\boldsymbol{y}(t))$ the equilibrium instantaneous stock market volatility, as defined in Eq.(10), the VIX index predicted by the model is

$$\text{VIX}(\boldsymbol{y}(t)) \equiv \sqrt{\frac{1}{T-t} \int_t^T \mathbb{E}[W(\boldsymbol{y}(u)) \mid \boldsymbol{y}(t)] \mathrm{d}u}, \tag{27}$$

15

where $\mathbb{E}$ is the expectation under the risk-neutral probability. The problem is that we do not know VIX $(\boldsymbol{y}(t))$ in closed-form. However, we can make a functional expansion of $\mathbb{E}[\mathrm{W}(\boldsymbol{y}(u)) \mid \boldsymbol{y}(t)]$, as follows,

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

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

$$\mathrm{VIX}(\boldsymbol{y}(t)) = \sqrt{\mathrm{W}(\boldsymbol{y}(t)) + \frac{1}{2} \mathcal{A} \mathrm{W}(\boldsymbol{y}(t))(T-t)} \tag{28}$$

where

$$\mathcal{A} \mathrm{W}(\boldsymbol{y}) = \nabla_{\boldsymbol{y}} \mathrm{W}(\boldsymbol{y})^{\mathsf{T}}(\boldsymbol{c} - \boldsymbol{D}\boldsymbol{y}) + \frac{1}{2} \left(\sum_{j=1}^{3} (\alpha_j + \beta_j y_j) \nabla_{y_j y_j} \mathrm{W}(\boldsymbol{y}) \right), \tag{29}$$

$$\mathrm{W}(\boldsymbol{y}(t)) = \frac{\sum_{j=1}^{3} s_j^2 (\alpha_j + \beta_j y_j(t))}{s(\boldsymbol{y}(t))^2} \tag{30}$$

$$\nabla_{y_j} \mathrm{W}(\boldsymbol{y}(t)) = \frac{s_j^2 \beta_j - 2\mathrm{W}(\boldsymbol{y}(t)) s(\boldsymbol{y}(t)) s_j}{s(\boldsymbol{y}(t))^2} \tag{31}$$

$$\nabla_{y_j y_j} \mathrm{W}(\boldsymbol{y}(t)) = -2 \frac{s_j}{s(\boldsymbol{y}(t))^2} \left(\frac{s_j^2 \beta_j}{s(\boldsymbol{y}(t))} + s(\boldsymbol{y}(t)) \mathrm{W}_{y_j}(\boldsymbol{y}(t)) - s_j \mathrm{W}(\boldsymbol{y}(t)) \right) \tag{32}$$

$$\boldsymbol{c} = \begin{bmatrix} \kappa_1 \mu_1 + \bar{\kappa}_1 \mu_2 - \alpha_1 \lambda_{1(1)} \\ \bar{\kappa}_2 \mu_1 + \kappa_2 \mu_2 - \alpha_2 \lambda_{1(2)} \\ \kappa_3 \mu_3 - \alpha_3 \lambda_{1(3)} \end{bmatrix}$$

$$\boldsymbol{D} = \begin{bmatrix} \kappa_1 + \lambda_{1(1)} \beta_1 + \lambda_{2(1)} & \bar{\kappa}_1 & 0 \\ \bar{\kappa}_2 & \kappa_2 + \lambda_{1(2)} \beta_2 + \lambda_{2(2)} & 0 \\ 0 & 0 & \kappa_3 + \lambda_{1(3)} \beta_3 + \lambda_{2(3)} \end{bmatrix} \tag{33}$$

In the actual computation of (29), (30), (31), (32), (33) we replace the unknown parameters $s_0, s_j, \kappa_j, \alpha_j, \beta_j \ j = 1, 2, 3$ and $\bar{\kappa}_i, \mu_i, \ i = 1, 2$ with their estimated counterparts computed in the previous two stages, i.e. using $\hat{\theta}_T, \hat{\phi}_T$. Also, in the construction of (30), (31), (32) we make use of actual samples for the observable factors $y_{1,t}, y_{2,t}$ and simulated samples for the latent factors, where the latter is simulated using the parameters estimated n the second step. Note that, given $\hat{\theta}_T, \hat{\phi}_T$, we can identify $\lambda = \left(\lambda_{1(1)}, \lambda_{1(2)}, \lambda_{1(3)}, \lambda_{2(1)}, \lambda_{2(2)}, \lambda_{2(3)} \right)^{\mathsf{T}}$ from $\boldsymbol{c}$ and $\boldsymbol{D}$ in (33). Let $\mathrm{VIX}_{t,h}^{\Delta} \left(\hat{\theta}_T, \hat{\phi}_T, \lambda \right)$ and VIX_t be the model-based and model free series of the VIX index. As the CBOE VIX index is available only since 1990 (?), in this stage we use a sample of length $\Upsilon < T$. In the sequel, we rely on the following auxiliary model

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

16

and on its simulation-based counterpart,

$$\text{VIX}_{t,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right) = a_h^{\text{VIX}} + \varphi_h \text{VIX}_{t-1,h}^{\Delta}\left(\hat{\theta}_T^{\Delta}, \hat{\phi}_T^{\Delta}, \lambda\right) + \sum_{i \in \{36,48\}} b_{1,i,h}^{\text{VIX}} y_{1,t-i} + \sum_{i \in \{36,48\}} b_{2,i,h}^{\text{VIX}} y_{2,t-i} + \epsilon_{h,t}^{\text{VIX}}.$$

Let $\mathbf{VY}_t = (y_{1,t-36}, y_{1,t-48}, y_{2,t-36}, y_{2,t-48}, \text{VIX}_{t-1})^{\mathsf{T}}$, and
$\mathbf{VY}_{t,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right) = \left(y_{1,t-36}, y_{1,t-48}, y_{2,t-36}, y_{2,t-48}, \text{VIX}_{t-1,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right)\right)^{\mathsf{T}}$, and define

$$\tilde{\psi}_{1,\Upsilon} = \left(\frac{1}{\Upsilon} \sum_{t=49}^{\Upsilon} \mathbf{VY}_t \mathbf{VY}_t^{\mathsf{T}}\right)^{-1} \frac{1}{\Upsilon} \sum_{t=49}^{\Upsilon} \mathbf{VY}_t \text{VIX}_t$$

$$\hat{\psi}_{1,\Upsilon,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right) = \left(\frac{1}{\Upsilon} \sum_{t=49}^{\Upsilon} \mathbf{VY}_{t,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right) \mathbf{VY}_{t,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right)^{\mathsf{T}}\right)^{-1}$$
$$\times \frac{1}{\Upsilon} \sum_{t=49}^{\Upsilon} \mathbf{VY}_{t,h}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right) \text{VIX}_{t,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right)$$

and finally define

$$\tilde{\psi}_{\Upsilon} = \left(\tilde{\psi}_{1,\Upsilon}, \overline{\text{VIX}}, \hat{\sigma}_{\text{VIX}}^2\right)^{\mathsf{T}}$$

$$\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right) = \left(\hat{\psi}_{1,\Upsilon,h}^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right), \overline{\text{VIX}}_h^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right), \hat{\sigma}_{\text{VIX}}^{2,\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right)\right)^{\mathsf{T}},$$

where $\overline{\text{VIX}}, \hat{\sigma}_{\text{VIX}}^2, \overline{\text{VIX}}_h^{\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right), \hat{\sigma}_{\text{VIX}}^{2,\Delta}\left(\hat{\theta}_T, \hat{\phi}_T, \lambda\right)$ denote sample mean and sample variance of model-free and model-based VIX.

The estimator for the risk premia parameters and its probability llimit are given by

$$\hat{\lambda}_{\Upsilon} = \arg\min_{\lambda \in \Lambda} \left(\left(\frac{1}{H} \sum_{h=1}^{H} \hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \lambda\right) - \tilde{\psi}_{\Upsilon}\right)^{\mathsf{T}} \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \lambda\right) - \tilde{\psi}_T\right)\right),$$

and

$$\lambda_0 = \arg\min_{\lambda \in \Lambda} \text{p} \lim_{\Upsilon \to \infty, \Delta \to 0} \left(\left(\frac{1}{H} \sum_{h=1}^{H} \hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \lambda\right) - \tilde{\psi}_{\Upsilon}\right)^{\mathsf{T}} \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \lambda\right) - \tilde{\psi}_{\Upsilon}\right)\right).$$

Proposition 4: *If as* $T, \Upsilon \to \infty$, $\Delta\sqrt{T} \to 0$, $\Delta T \to \infty$, $\Upsilon/T \to \pi$, $0 < \pi < 1$,

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

where

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

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

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

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

$$= \operatorname{Acov}\left(\sqrt{\Upsilon} \left(\hat{\psi}_{\Upsilon,h'}^\Delta \left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0 \right), \sqrt{\Upsilon} \left(\hat{\psi}_{\Upsilon,h}^\Delta \left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0 \right)^\intercal \right), \ \forall \ h \neq h'$$

and

$$\mathbf{P}_3 = \pi F_{\theta_0}^\intercal \operatorname{Avar}\left(\sqrt{T} \left(\hat{\theta}_T - \theta_0 \right) \right) F_{\theta_0} + \pi F_{\phi_0}^\intercal \operatorname{Avar} \sqrt{T} \left(\hat{\phi}_T - \phi_0 \right) F_{\phi_0}$$

$$+ 2\pi \operatorname{Acov}\left(F_{\phi_0}^\intercal \sqrt{T} \left(\hat{\phi}_T - \phi_0 \right), F_{\theta_0}^\intercal \sqrt{T} \left(\hat{\theta}_T - \theta_0 \right) \right)$$

$$+ 2\sqrt{\pi} \operatorname{Acov}\left(\sqrt{T} \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\psi}_{\Upsilon,h}^\Delta \left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0 \right), F_{\phi_0}^\intercal \sqrt{T} \left(\hat{\phi}_T - \phi_0 \right) \right)$$

$$+ 2\sqrt{\pi} \operatorname{Acov}\left(\sqrt{T} \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\psi}_{\Upsilon,h}^\Delta \left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0 \right), F_{\theta_0}^\intercal \sqrt{T} \left(\hat{\theta}_T - \theta_0 \right) \right)$$

$$- 2\sqrt{\pi} \operatorname{Acov}\left(\sqrt{T} \left(\sqrt{\Upsilon} \left(\tilde{\psi}_\Upsilon - \boldsymbol{\psi}_0 \right) \right), F_{\phi_0}^\intercal \sqrt{T} \left(\hat{\phi}_T - \phi_0 \right) \right)$$

$$- 2\sqrt{\pi} \operatorname{Acov}\left(\sqrt{T} \left(\sqrt{\Upsilon} \left(\tilde{\psi}_\Upsilon - \boldsymbol{\psi}_0 \right) \right), F_{\theta_0}^\intercal \sqrt{T} \left(\hat{\theta}_T - \theta_0 \right) \right)$$

with $F_{\theta_0}^\intercal = \operatorname{p\,lim}_{T,\Upsilon \to \infty} \nabla_\theta \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\psi}_{\Upsilon,h} \left(\phi_0, \theta_0, \lambda_0\right) \right),$
$F_{\phi_0}^\intercal = \operatorname{p\,lim}_{T,\Upsilon \to \infty} \nabla_\phi \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\psi}_{\Upsilon,h} \left(\phi_0, \theta_0, \lambda_0\right) \right), \ and \ \Upsilon/T \to \pi.$

Note that the matrix $\mathbf{P}_3$ captures the contribution of parameter estimation error, due to the fact that the model VIX series has been simulated using parameters estimated in the previous two stages, i.e. $\hat{\phi}_T$ 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

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

the set of resampled observations.

3.4.1 Bootstrap Standard Error for ϕ

The simulated samples for $y_{1,t}$ and $y_{2,t}$ are independent of the actual samples and also are independent across simulation replications. Also, as stated in Lemma 1, the estimators of the auxiliary model parameters, based on actual and simulated samples, have the same asymptotic variance. Hence, there is no need to resample the simulated series. 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 recentering term. Broadly speaking, in the overidentified case, even if the population moment conditions have mean zero, the bootstrap moment conditions do not have mean zero, and a proper recentering term is necessary (see e.g. Hall and Horowitz 1996).

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

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

where $\tilde{\varphi}_{1,T}^*, \tilde{\varphi}_{2,T}^*$ are the estimated parameters of the auxiliary models computed using resampled observations,[3] and $\bar{y}_1^*, \bar{y}_2^*, \hat{\sigma}_1^{*2}, \hat{\sigma}_2^{*2}$ are sample mean and variance of $y_{1,t}^*, y_{2,t}^*$. Define,

$$
\begin{aligned}
\hat{\phi}_T^* &= \arg\min_{\varphi} \left(\left(\left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\varphi}_{T,h}^\Delta(\phi) - \hat{\varphi}_{T,h}^\Delta\left(\hat{\phi}_T\right) \right) \right) - \left(\tilde{\varphi}_T^* - \tilde{\varphi}_T\right) \right)^\top \right. \\
&\qquad \left. \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\varphi}_{T,h}^\Delta(\phi) - \hat{\varphi}_{T,h}^\Delta\left(\hat{\phi}_T\right) \right) \right) - \left(\tilde{\varphi}_T^* - \tilde{\varphi}_T\right) \right)
\end{aligned}
$$

---

[3]For example, for $i = 1, 2$

$$\widetilde{\varphi}_{i,T}^* = \left(\frac{1}{T} \sum_{t=25}^{T} \mathbf{Y}_t^* \mathbf{Y}_t^{*\top} \right)^{-1} \frac{1}{T} \sum_{t=25}^{T} \mathbf{Y}_t^* y_{i,t}^*$$

We now construct B bootstrap estimators $\hat{\phi}^*_{T,i}$, and we construct the bootstrap covariance matrix, as

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

We then obtain asymptotically valid bootstrap standard errors from $\left(1 + \frac{1}{H}\right)\hat{\mathbf{V}}^*_{\phi_0,T,B}$.

Proposition 5: *Under the conditions in Lemma 1, if as $T, B \to \infty$, $l/T^{1/2} \to 0$,*

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

3.4.2   Bootstrap Standard Error for θ

The (model based) stock price series has been generated using actual samples for 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 resampled simulated stock price series as:

$$s^{*\Delta}_{t,h}(\theta) = s_0 + s_1 y^*_{1,t} + s_2 y^*_{2,t} + Z^{*\Delta}_{t,h}(\theta_u), \tag{34}$$

where $Z^{*\Delta}_{t,h}(\theta_u)$ is resampled from the simulated unobservable process $Z^{\Delta}_{t,h}(\theta_u)$, and use $s^{*\Delta}_{t,h}(\theta)$ to construct $R^{*\Delta}_{t,h}(\theta)$ and $\text{Vol}^{*\Delta}_{t,h}(\theta)$. Define,

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

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

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

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

Define:[4]

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

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

Construct the bootstrap covariance matrix, as

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

We have:

Proposition 6: *Under the conditions in Lemma 2, if as $T, B \to \infty$, $l/T^{1/2} \to 0$,*

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

3.4.3 Bootstrap Standard Error 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 Υ, instead of length T. Thus, we need to resample $y_{1,t}, y_{2,t}, s_t$ and VIX_t from the shorter sample, using blocksize l_Υ and number of blocks b_Υ, so that $l_\Upsilon b_\Upsilon = \Upsilon$. Also, we need to resample the unobservable factor $Z_{t,h}^\Delta \left(\hat{\theta}_T \right)$ from a sample of length Υ. Let $\text{VIX}_{t,h}^{*\Delta} \left(\hat{\phi}_T^*, \hat{\theta}_T^*, \lambda \right)$ be the model-based VIX index constructed using $y_{1,t}^*, y_{2,t}^*, Z_{t,h}^{*\Delta} \left(\hat{\theta}_T \right)$ and the bootstrap estimators $\hat{\phi}_T^*$ and $\hat{\theta}_T^*$. Finally, let

$$\tilde{\psi}_\Upsilon^* = \left(\tilde{\psi}_{1,\Upsilon}^*, \overline{\text{VIX}}^*, \hat{\sigma}_{\text{VIX}}^{2*} \right)^\intercal,$$

where $\tilde{\psi}_{1,\Upsilon}^*$ are the auxiliary model parameters estimated using $y_{1,t}^*, y_{2,t}^*$, and VIX_t^*, with VIX_t^* being the resampled series of the model free VIX, and $\overline{\text{VIX}}^*, \hat{\sigma}_{\text{VIX}}^{2*}$ are the sample mean and variance of VIX_t^*, and[5]

$$\hat{\psi}_{\Upsilon,h}^{*\Delta} \left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda \right) = \left(\hat{\psi}_{1,\Upsilon,h}^* \left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda \right), \overline{\text{VIX}}_h^{*\Delta} \left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda \right), \tilde{\sigma}_{\text{VIX}}^{*2,\Delta} \left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda \right) \right)^\intercal,$$

---

[4] For example,

$$\tilde{\vartheta}_{1,T}^* = \left(\frac{1}{T} \sum_{t=13}^T \mathbf{Y}_t^* \mathbf{Y}_t^{*\intercal} \right)^{-1} \frac{1}{T} \sum_{t=13}^T \mathbf{Y}_t^* R_t^*$$

and

$$\hat{\vartheta}_{1,T,h}^{*\Delta}(\theta) = \left(\frac{1}{T} \sum_{t=13}^T \mathbf{Y}_t^* \mathbf{Y}_t^{*\intercal} \right)^{-1} \frac{1}{T} \sum_{t=13}^T \mathbf{Y}_t^* R_{t,h}^{*\Delta}(\theta)$$

[5] For example,

$$\tilde{\psi}_{1,\Upsilon}^* = \left(\frac{1}{\Upsilon} \sum_{t=49}^\Upsilon \mathbf{V} \mathbf{Y}_t^* \mathbf{V} \mathbf{Y}_t^{*\intercal} \right)^{-1} \frac{1}{\Upsilon} \sum_{t=49}^\Upsilon \mathbf{V} \mathbf{Y}_t^* \text{VIX}_t^*$$

where $\hat{\psi}_{1,\Upsilon,h}^{*\Delta}\left(\hat{\phi}_T^*, \hat{\theta}_T^*, \lambda\right)$ are the auxiliary model parameters estimated using $y_{1,t}^*, y_{2,t}^*$, $\mathrm{VIX}_{t,h}^{*\Delta}\left(\hat{\phi}_T^*, \hat{\theta}_T^*, \lambda\right)$ and $\overline{\mathrm{VIX}}_h^{*\Delta}\left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda\right), \widetilde{\sigma}_{\mathrm{VIX}}^{*2,\Delta}\left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda\right)$ are the sample mean and variance of $\mathrm{VIX}_{t,h}^{*\Delta}\left(\hat{\phi}_T^*, \hat{\theta}_T^*, \lambda\right)$. Define,

$$
\begin{aligned}
\hat{\lambda}_\Upsilon^* \;=\; \arg\min_\lambda &\left(\left(\frac{1}{H}\sum_{h=1}^H\left(\hat{\psi}_{\Upsilon,h}^{*\Delta}\left(\hat{\phi}_T^*, \hat{\theta}_T^*, \lambda\right) - \hat{\psi}_{\Upsilon,h}^\Delta\left(\hat{\phi}_T, \hat{\theta}_T, \hat{\lambda}_\Upsilon\right)\right) - \left(\tilde{\boldsymbol{\psi}}_\Upsilon^* - \tilde{\boldsymbol{\psi}}_\Upsilon\right)\right)^\intercal \right.\\
&\left. \left(\frac{1}{H}\sum_{h=1}^H\left(\hat{\psi}_{\Upsilon,h}^{*\Delta}\left(\hat{\phi}_T^*, \hat{\theta}_T^*, \lambda\right) - \hat{\psi}_{\Upsilon,h}^\Delta\left(\hat{\phi}_T, \hat{\theta}_T, \hat{\lambda}_\Upsilon\right)\right) - \left(\tilde{\boldsymbol{\psi}}_\Upsilon^* - \tilde{\boldsymbol{\psi}}_\Upsilon\right)\right)\right)
\end{aligned}
$$

Construct the bootstrap covariance matrix, as

$$
\hat{\mathbf{V}}_{\lambda_0,\Upsilon,B}^* = \Upsilon\frac{1}{B}\sum_{i=1}^B\left(\left(\hat{\lambda}_{\Upsilon,i}^* - \frac{1}{B}\sum_{i=1}^B\hat{\lambda}_{\Upsilon,i}^*\right)\left(\hat{\lambda}_{\Upsilon,i}^* - \frac{1}{B}\sum_{i=1}^B\hat{\lambda}_{\Upsilon,i}^*\right)^\intercal\right).
$$

We have:

Proposition 7: *Under the conditions in Lemma 3, if as $T, \Upsilon, B \to \infty$, $l_\Upsilon/T_\Upsilon^{1/2} \to 0$,*

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

4 Empirical analysis

4.1 Data

Our sample data 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,

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

where CPI_t is the consumer price index and IP_t is the seasonally adjusted industrial production index, as of month t. Figure 1 depicts the two series $y_{1,t}$ and $y_{2,t}$, along with NBER-dated recession events. As for the stock price data, we use the S&P Compounded index and the VIX index. 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 23rd of every month. Thus, to possibly avoid overreaction to releases of information, we sample the S&P Compounded index and the VIX index every 25th of the month.

---

and

$$
\hat{\psi}_{1,\Upsilon,h}^{\Delta*}\left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda\right) = \left(\frac{1}{\Upsilon}\sum_{t=49}^\Upsilon \mathbf{VY}_{t,h}^{*\Delta}\left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda\right)\mathbf{VY}_{t,h}^{*\Delta}\left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda\right)^\intercal\right)^{-1}\frac{1}{\Upsilon}\sum_{t=49}^\Upsilon \mathbf{VY}_{t,h}^{*\Delta}\left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda\right)\mathrm{VIX}_{t,h}^{*\Delta}\left(\hat{\theta}_T^*, \hat{\phi}_T^*, \lambda\right)
$$

4.2   Estimation results

Tables 1 through 3 report parameter estimates.

Figure 2 depicts sample data related to the continuously compounded price changes, R_t, return volatility, Vol_t, along with the dynamics predicted by the model.

Figure 3 (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 bottom panel in Figure 3 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

Figure 4 provides scatterplots of the volatility risk-premium against inflation and industrial production.

5   Conclusion

This paper develops a framework to analyze the business cycle movements of stock market returns, volatility and volatility risk-premia. In our model, the aggregate stock market behavior relates to the development of two macroeconomic factors, inflation and industrial production growth, and one unobserved factor. The relations linking the asset price, returns and volatility to these factors are derived under 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.

We take our model to data, and make use of the new volatility index, the VIX index, to estimate the parameters related to risk-aversion. Our model predicts that stock market returns are procyclical, stock market volatility is countercyclical and volatility risk-premia are counter-cyclical.

Appendix

A. Proofs for Section 2

Existence of a strong solution to Eq. (1). 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 model we consider in the empirical section, for which $\boldsymbol{\Sigma} = \boldsymbol{I}_{3\times 3}$, $\boldsymbol{\beta}_i$ is a vector of zeros, except possibly for its i-th element, denoted as $\beta_i \equiv \beta_{ii}$, and $\boldsymbol{\kappa}$ is as in Section 2.5. Condition (i) collapses to,

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

$$\text{For all } y_3 \quad : \quad \alpha_3 + \beta_3 y_3 = 0, \quad \beta_3 \kappa_3 (\mu_3 - y_3) > \frac{1}{2}\beta_3^2$$

That is, ruling out the trivial case $\beta_i = 0$,

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

$$\kappa_3 (\mu_3 \beta_3 + \alpha_3) > \frac{1}{2}\beta_3^2$$

Proof of Proposition 1. 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,

$$\frac{\mathrm{d}s^\xi(t)}{s^\xi(t)} = \mathrm{D}(\boldsymbol{y}(t))\,\mathrm{d}t + \left(\mathbf{Q}(\boldsymbol{y}(t))^\top - \boldsymbol{\Lambda}(\boldsymbol{y}(t))^\top\right)\mathrm{d}\boldsymbol{W}(t), \tag{A1}$$

where

$$\mathrm{D}(\boldsymbol{y}) \quad \equiv \quad -r + \frac{\mathcal{A}s(\boldsymbol{y})}{s(\boldsymbol{y})} - \mathbf{Q}(\boldsymbol{y})^\top \boldsymbol{\Lambda}(\boldsymbol{y}) \quad ; \quad \mathbf{Q}(\boldsymbol{y})^\top = \frac{s_y(\boldsymbol{y})^\top \boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})}{s(\boldsymbol{y})}$$

$$\mathcal{A}s(\boldsymbol{y}) \quad \equiv \quad s_y(\boldsymbol{y})^\top \boldsymbol{\kappa}(\boldsymbol{\mu} - \boldsymbol{y}) + \frac{1}{2}\mathrm{Tr}\left([\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})][\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})]^\top s_{yy}(\boldsymbol{y})\right).$$

By absence of arbitrage opportunities, for any $T < \infty$,

$$s^\xi(t) = E\left[\int_t^T \delta^\xi(h)\,\mathrm{d}h \,\middle|\, F(t)\right] + E[s^\xi(T) \mid F(t)], \tag{A2}$$

24

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

$$\lim_{T\to\infty} E[s^\xi(T) \mid F(t)] = 0, \tag{A3}$$

from which Eq. (7) in the main text follows, once we show that $\int_t^\infty E\left[\delta^\xi(h)\right] dh < \infty$.

Next, by Eq. (A2),

$$0 = \frac{d}{d\tau} E[s^\xi(\tau) \mid F(t)]\bigg|_{\tau=t} + \delta^\xi(t). \tag{A4}$$

Below, we show that

$$E[s^\xi(T) \mid F(t)] = s^\xi(t) + \int_t^T \mathrm{D}\left(\boldsymbol{y}(h)\right) s^\xi(h)\, dh. \tag{A5}$$

Therefore, by the assumptions on $\boldsymbol{\Lambda}$, Eq. (A4) can be rearranged to yield the following partial differential equation,

$$\text{For all } \boldsymbol{y}, \quad s_y(\boldsymbol{y})^\top (\boldsymbol{c} - \boldsymbol{D}\boldsymbol{y}) + \frac{1}{2}\mathrm{Tr}\left([\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})][\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})]^\top s_{yy}(\boldsymbol{y})\right) + \delta(\boldsymbol{y}) - rs(\boldsymbol{y}) = 0, \tag{A6}$$

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

Let us assume that the price function is affine in $\boldsymbol{y}$,

$$s(\boldsymbol{y}) = \gamma + \boldsymbol{\eta}^\top \boldsymbol{y}, \tag{A7}$$

for some scalar γ and some vector $\boldsymbol{\eta}$. By plugging this guess back into Eq. (A6) 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,

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

We are left to show that Eq. (A3) and (A5) hold true.

As regards Eq. (A3), we have

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

where the second line follows by Eq. (A7), and the third line holds because $E[\xi(T) \mid \mathbb{F}(t)] = 1$ for all T, and by a change of measure (**Warning**: We need some more work to prove this.) Eq. (A3) follows because $\boldsymbol{y}$ is stationary mean-reverting.

To show that Eq. (A5) holds, we need to show that the diffusion part of s^ξ in Eq. (A1) is a martingale, not only a local martingale, which it does whenever for all T,

$$
\infty > E\left[\int_t^T \left\|\boldsymbol{Q}\left(\boldsymbol{y}\left(\tau\right)\right)^\top - \boldsymbol{\Lambda}\left(\tau\right)^\top\right\|^2 \mathrm{d}\tau\right] = E\left[\int_t^T \left\|\frac{\boldsymbol{\eta}^\top \boldsymbol{\Sigma} \boldsymbol{V}\left(\boldsymbol{y}\left(\tau\right)\right)}{\gamma + \boldsymbol{\eta}^\top \boldsymbol{y}\left(\tau\right)} - \boldsymbol{\Lambda}\left(\tau\right)^\top\right\|^2 \mathrm{d}\tau\right].
$$

∎

B. Proofs for Section 3

Proof of Proposition 2: By the first order conditions and a mean value expansion around ϕ_0,

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

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

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

Let $\hat{\varphi}_{T,h}\left(\phi\right)$ be the estimator obtained in the case we simulated continuous paths of $y_{1,t,h}\left(\phi\right), y_{2,t,h}\left(\phi\right),$ i.e. $\Delta = 0$. As $\Delta\sqrt{T} \to 0$, by Pardoux and Talay (1985),

$$
\mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H \hat{\varphi}_{T,h}^\Delta\left(\phi_0\right) - \varphi_0\right)\right) = \mathrm{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^H \hat{\varphi}_{T,h}\left(\phi_0\right) - \varphi_0\right)\right)
$$

Thus,

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

26

Now,

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

$$
=\quad \text{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}\left(\phi_0\right)-\varphi_0\right)\right)+\text{Avar}\left(\sqrt{T}\left(\tilde{\varphi}_T-\varphi_0\right)\right)
$$

$$
-2\text{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}\left(\phi_0\right)-\varphi_0\right),\sqrt{T}\left(\tilde{\varphi}_T-\varphi_0\right)\right).
$$

As the simulated paths are independent of the sample paths,

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

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

$$
\text{Avar}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\varphi}_{T,h}\left(\phi_0\right)-\varphi_0\right)\right)
$$

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

$$
=\frac{1}{H}\text{Avar}\sqrt{T}\left(\hat{\varphi}_{T,1}\left(\phi_0\right)-\varphi_0\right)=\frac{1}{H}\mathbf{J}_1.
$$

Finally, $\text{Avar}\left(\sqrt{T}\left(\tilde{\varphi}_T-\varphi_0\right)\right)=\text{Avar}\sqrt{T}\left(\hat{\varphi}_{T,1}\left(\phi_0\right)-\varphi_0\right)=\mathbf{J}_1$, 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 statement in the Lemma then follows by the central limit theorem for geometrically strong mixing processes.

Proof of Proposition 3: By the same argument as 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,

$$\mathrm{Avar}\left(\sqrt{T}\left(\hat{\theta}_T - \theta_0\right)\right)$$

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

$$\left. -2\mathrm{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}^{\Delta}(\theta_0) - \vartheta_0\right), \sqrt{T}\left(\tilde{\vartheta}_T - \vartheta_0\right)\right)\right)\mathbf{D}_2\left(\mathbf{D}_2^\intercal \mathbf{D}_2\right)^{-1}.$$

Let $\hat{\vartheta}_{T,h}(\theta_0)$ be the estimator obtained in the case we simulated continuous paths for the unobservable factor $Z_{t,h}(\theta)$, i.e. $\Delta = 0$, then by the same argument as in the proof of Proposition 2,

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

In the current context, paths for the model-based stock return have been simulated using sample paths for the observable factors $y_{1,t}, y_{2,t}$, hence simulated paths are no longer independent across simulation replications, and are no longer independent of the actual sample paths of stock return and volatility. Thus,

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

$$= \frac{1}{H}\mathrm{Avar}\sqrt{T}\left(\hat{\vartheta}_{T,1}(\theta_0) - \vartheta_0\right) + \frac{H(H-1)}{H^2}\mathrm{Acov}\left(\sqrt{T}\left(\hat{\vartheta}_{T,1}(\theta_0) - \vartheta_0\right), \sqrt{T}\left(\hat{\vartheta}_{T,h}(\theta_0) - \vartheta_0\right)\right)$$

$$= \frac{1}{H}\mathbf{J}_2 + \frac{H(H-1)}{H^2}\mathbf{K}_2,$$

and

$$\mathrm{Acov}\left(\sqrt{T}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\vartheta}_{T,h}(\theta_0) - \vartheta_0\right), \sqrt{T}\left(\tilde{\vartheta}_T - \vartheta_0\right)\right)$$

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

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

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

Proof of Proposition 4:

28

Given the first order conditions, and by a mean value expansion around λ_0,

$$
\sqrt{\Upsilon}\left(\hat{\lambda}_{\Upsilon} - \lambda_0\right)
$$

$$
= -\left(\nabla_\lambda\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \hat{\lambda}_{\Upsilon}\right)\right)^{\intercal}\nabla_\lambda\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \bar{\lambda}_{\Upsilon}\right)\right)\right)^{-1}
$$

$$
\nabla_\lambda\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \hat{\lambda}_{\Upsilon}\right)\right)^{\intercal}\sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \lambda_0\right) - \tilde{\boldsymbol{\psi}}_{\Upsilon}\right),
$$

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

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

Now,

$$
\sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \lambda_0\right) - \boldsymbol{\psi}_0\right)
$$

$$
= \sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right) + \sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \hat{\theta}_T, \lambda_0\right) - \hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \theta_0, \lambda_0\right)\right)\right)
$$

$$
+ \sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \theta_0, \lambda_0\right) - \hat{\psi}_{\Upsilon,h}^{\Delta}\left(\phi_0, \theta_0, \lambda_0\right)\right)\right)
$$

$$
= \sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right) + \nabla_\theta\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\hat{\phi}_T, \bar{\theta}_T, \lambda_0\right)\right)\sqrt{\Upsilon}\left(\hat{\theta}_T - \theta_0\right)
$$

$$
+ \nabla_\phi\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\bar{\phi}_T, \theta_0, \lambda_0\right)\right)\sqrt{\Upsilon}\left(\hat{\phi}_T - \phi_0\right)
$$

$$
= \sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^{H}\hat{\psi}_{\Upsilon,h}^{\Delta}\left(\phi_0, \theta_0, \lambda_0\right) - \boldsymbol{\psi}_0\right) + F_{\theta_0}^{\intercal}\sqrt{\pi}\sqrt{T}\left(\hat{\theta}_T - \theta_0\right) + F_{\phi_0}^{\intercal}\sqrt{\pi}\sqrt{T}\left(\hat{\phi}_T - \phi_0\right) + o_p(1),
$$

where $\bar{\theta}_T \in \left(\hat{\theta}_T, \theta_0\right), \bar{\phi}_T \in \left(\hat{\phi}_T, \phi\right) \pi = \lim_{T, \Upsilon \to \infty} \Upsilon/T$, and

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

Now,

$$
\begin{aligned}
&\mathrm{Avar}\sqrt{\Upsilon}\left(\hat{\lambda}_\Upsilon - \lambda_0\right) \\
=&\ (\mathbf{D}_3^\intercal \mathbf{D}_3)^{-1}\mathbf{D}_3^\intercal\left(\mathrm{Avar}\sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^H \hat{\psi}_{\Upsilon,h}^\Delta(\phi_0,\theta_0,\lambda_0) - \boldsymbol{\psi}_0\right) + \mathrm{Avar}\sqrt{\Upsilon}\left(\tilde{\boldsymbol{\psi}}_\Upsilon - \boldsymbol{\psi}_0\right)\right. \\
&-2\mathrm{Acov}\left(\sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^H \hat{\psi}_{\Upsilon,h}^\Delta(\phi_0,\theta_0,\lambda_0) - \boldsymbol{\psi}_0\right),\sqrt{\Upsilon}\left(\tilde{\boldsymbol{\psi}}_\Upsilon - \boldsymbol{\psi}_0\right)\right) + \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}\sqrt{T}\left(\hat{\phi}_T - \phi_0\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}_{\Upsilon,h}^\Delta(\phi_0,\theta_0,\lambda_0) - \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}_{\Upsilon,h}^\Delta(\phi_0,\theta_0,\lambda_0) - \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{\Upsilon}\left(\tilde{\boldsymbol{\psi}}_\Upsilon - \boldsymbol{\psi}_0\right)\right),F_{\phi_0}^\intercal\sqrt{T}\left(\hat{\phi}_T - \phi_0\right)\right) \\
&\left.-2\sqrt{\pi}\mathrm{Acov}\left(\sqrt{T}\left(\sqrt{\Upsilon}\left(\tilde{\boldsymbol{\psi}}_\Upsilon - \boldsymbol{\psi}_0\right)\right),F_{\theta_0}^\intercal\sqrt{T}\left(\hat{\theta}_T - \theta_0\right)\right)\right)\mathbf{D}_3\left(\mathbf{D}_3^\intercal\mathbf{D}_3\right) \quad (35)
\end{aligned}
$$

Let $\hat{\psi}_{\Upsilon,h}\left(\hat{\phi}_T,\hat{\theta}_T,\lambda_0\right)$ 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,

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

By the same argument as in Proposition 3,

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

Hence given Propositions 2 and 3,

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

30

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 (35).

Hereafter, let P^* be the probability measure governing the resampled series, and $\mathrm{E}^*, \mathrm{var}^*$ 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.

Proof of Proposition 5: By the first order conditions and a mean value expansion around $\hat{\phi}_T$,

$$
0 = \nabla_\phi \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\varphi}_{T,h}^\Delta \left(\hat{\phi}_T^* \right) \right)^\mathsf{T} \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\varphi}_{T,h}^\Delta \left(\hat{\phi}_T^* \right) - \hat{\varphi}_{T,h}^\Delta \left(\hat{\phi}_T \right) \right) - \left(\tilde{\varphi}_T^* - \tilde{\varphi}_T \right) \right)
$$

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

where $\bar{\phi}_T^* \in \left(\hat{\phi}_T^*, \hat{\phi}_T \right)$. Hence,

$$
\sqrt{T} \left(\hat{\phi}_T^* - \hat{\phi}_T \right)
$$

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

The statement in the Proposition will follow once we have shown that:

$$
\mathrm{E}^* \left(\sqrt{T} \left(\tilde{\varphi}_T^* - \tilde{\varphi}_T \right) \right) = o_p(1) \tag{36}
$$

$$
\mathrm{var}^* \left(\sqrt{T} \left(\tilde{\varphi}_T^* - \tilde{\varphi}_T \right) \right) = \mathrm{var} \left(\sqrt{T} \left(\tilde{\varphi}_T - \varphi_0 \right) \right) + O_p(l/\sqrt{T}) \tag{37}
$$

and for $\delta > 0$,

$$
\mathrm{E}^* \left(\left(\sqrt{T} \left(\tilde{\varphi}_T^* - \tilde{\varphi}_T \right) \right)^{2+\delta} \right) = O_p(1) \tag{38}
$$

In fact, given (36) and (37), by the uniform law of large numbers, $\nabla_\phi \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\varphi}_{T,h} \left(\hat{\phi}_T^* \right) \right) - \mathbf{D}_1 = o_p^*(1)$. Hence,

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

and, given (37), 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 (38), the statement then follows from Theorem 1 in Goncalves and White (2005).

It remains to show (36),(37) and (38). Now,

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

and as all components of $\sqrt{T}\left(\tilde{\varphi}_T^* - \tilde{\varphi}_T\right)$ can be treated in an analogous manner, we just consider $\sqrt{T}\left(\tilde{\varphi}_{1,T}^* - \tilde{\varphi}_{1,T}\right)$. By the first order conditions,

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

as $\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_t^*\mathbf{Y}_t^{*\intercal} - \mathrm{E}^*\left(\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_t^*\mathbf{Y}_t^{*\intercal}\right) = o_p^*(1)$, and $\mathrm{E}^*\left(\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_t^*\mathbf{Y}_t^{*\intercal}\right) = \frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_t\mathbf{Y}_t^{\intercal} + O_p(l/T) = \mathrm{E}\left(\mathbf{Y}_t\mathbf{Y}_t^{\intercal}\right) + o_p(1)$. Recalling (22),

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

This proves (36). Now,

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

where $\widetilde{\epsilon}_{1,t} = y_{1,t} - \mathbf{Y}_t^T\tilde{\varphi}_{1,T}$. This proves (37). Finally, as $\frac{1}{T}\sum_{t=25}^{T}\mathbf{Y}_t\mathbf{Y}_t$ is full rank, by the same argument as above,

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

This proves (38).

Proof of Proposition 6: By the first order conditions and a mean value expansion around $\hat{\theta}_T$,

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

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

$$
+ \nabla_\theta \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{*\Delta} \left(\hat{\theta}_T^* \right) \right)^\intercal \nabla_\theta \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{*\Delta} \left(\bar{\theta}_T^* \right) \right) \left(\hat{\theta}_T^* - \hat{\theta}_T \right),
$$

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

$$
\sqrt{T} \left(\hat{\theta}_T^* - \hat{\theta}_T \right)
$$

$$
= -\left(\nabla_\theta \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{*\Delta} \left(\hat{\theta}_T^* \right) \right)^\intercal \nabla_\theta \left(\frac{1}{H} \sum_{h=1}^{H} \hat{\vartheta}_{T,h}^{*\Delta} \left(\bar{\theta}_T^* \right) \right) \right)^{-1}
$$

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

We need to show that:

$$
\mathrm{E}^* \left(\sqrt{T} \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\vartheta}_{T,h}^{*\Delta} \left(\hat{\theta}_T \right) - \hat{\vartheta}_{T,h}^{\Delta} \left(\hat{\theta}_T \right) \right) \right) \right) = o_p(1) \tag{39}
$$

$$
\mathrm{var}^* \left(\sqrt{T} \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\vartheta}_{T,h}^{*\Delta} \left(\hat{\theta}_T \right) - \hat{\vartheta}_{T,h}^{\Delta} \left(\hat{\theta}_T \right) \right) \right) \right) = \mathrm{var} \left(\sqrt{T} \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\vartheta}_{T,h}^{\Delta} \left(\hat{\theta}_T \right) - \vartheta \left(\theta_0 \right) \right) \right) \right) + o_p(1) \tag{40}
$$

$$
\mathrm{E}^* \left(\left(\sqrt{T} \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\vartheta}_{T,h}^{*\Delta} \left(\hat{\theta}_T \right) - \hat{\vartheta}_{T,h}^{\Delta} \left(\hat{\theta}_T \right) \right) \right) \right)^{2+\delta} \right) < \infty. \tag{41}
$$

The statement in the Proposition will then follows by the same argument used in the proof of Proposition 5. Note that,

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

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

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

$$=\frac{1}{H}\sum_{h=1}^{H}\mathrm{E}^*\left(\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}_{t,h}^{*\Delta}\left(\hat{\theta}_T\right)\mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}_t^{*\Delta}\left(\hat{\theta}_T\right)\left(R_{t,h}^{*\Delta}(\hat{\theta}_T)-\mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\hat{\vartheta}_{1,T,h}^{\Delta}\left(\hat{\theta}_T\right)\right)\right)$$

$$=\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}_{t,h}^{\Delta}\left(\hat{\theta}_T\right)\mathbf{Y}_{t,h}^{\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\right)^{-1}\mathrm{E}^*\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}_t^{*\Delta}\left(\hat{\theta}_T\right)\left(R_{t,h}^{*\Delta}(\hat{\theta}_T)-\mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\hat{\vartheta}_{1,T,h}^{\Delta}\left(\hat{\theta}_T\right)\right)\right)$$

$$+\frac{1}{H}\sum_{h=1}^{H}\mathrm{E}^*\left(\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}_{t,h}^{*\Delta}\left(\hat{\theta}_T\right)\mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\right)^{-1}-\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}_{t,h}^{\Delta}\left(\hat{\theta}_T\right)\mathbf{Y}_{t,h}^{\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\right)^{-1}\right)$$

$$\times\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}_t^{*\Delta}\left(\hat{\theta}_T\right)\left(R_{t,h}^{*\Delta}(\hat{\theta}_T)-\mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\hat{\vartheta}_{1,T,h}^{\Delta}\left(\hat{\theta}_T\right)\right)\right)$$

$$=\mathrm{E}^*\left(I_{T,h}^*\right)+\mathrm{E}^*\left(II_{T,h}^*\right)$$

Recalling (26),

$$\mathrm{E}^*\left(I_{T,h}^*\right)\;=\;\frac{1}{H}\sum_{h=1}^{H}\left(\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}_{t,h}^{\Delta}\left(\hat{\theta}_T\right)\mathbf{Y}_{t,h}^{\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\right)^{-1}\right.$$

$$\left.\times\,\mathrm{E}^*\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}_t^{*\Delta}\left(\hat{\theta}_T\right)\left(R_{t,h}^{*\Delta}(\hat{\theta}_T)-\mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\hat{\vartheta}_{1,T,h}^{\Delta}\left(\hat{\theta}_T\right)\right)\right)\right)$$

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

and as $II_{T,h}^*$ is of smaller order that $I_{T,h}^*$, $\mathrm{E}^*\left(II_{T,h}^*\right)=o_p(1)$. This proves (39). As for $h=1,...,H$

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

$$=\mathrm{E}\left(\left(\frac{1}{T}\sum_{t=13}^{T}\mathbf{Y}_{t,h}^{\Delta}\left(\theta_0\right)\mathbf{Y}_t^{\mathsf{T}\Delta}\left(\theta_0\right)\right)^{-1}\right)+o_p(1),$$

it suffices to show that

$$\mathrm{var}^*\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}_t^{*\Delta}\left(\hat{\theta}_T\right)\left(R_{t,h}^{*\Delta}(\hat{\theta}_T)-\mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\hat{\vartheta}_{1,T,h}^{\Delta}\left(\hat{\theta}_T\right)\right)\right)\right)$$

$$=\;\mathrm{Avar}\left(\frac{1}{H}\sum_{h=1}^{H}\left(\frac{1}{\sqrt{T}}\sum_{t=13}^{T}\mathbf{Y}_t^{\Delta}\left(\hat{\theta}_T\right)\left(R_{t,h}^{\Delta}(\hat{\theta}_T)-\mathbf{Y}_t^{\mathsf{T}\Delta}\left(\hat{\theta}_T\right)\hat{\vartheta}_{1,T,h}^{\Delta}\left(\hat{\theta}_T\right)\right)\right)\right)+o_p(1)$$

34

Now, recalling that the blocks are independent each other

$$\text{var}^* \left(\frac{1}{H} \sum_{h=1}^H \left(\frac{1}{\sqrt{T}} \sum_{t=13}^T \mathbf{Y}_t^{*\Delta}\left(\hat\theta_T\right) \left(R_{t,h}^{*\Delta}(\hat\theta_T) - \mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat\theta_T\right) \hat\vartheta_{1,T,h}^\Delta\left(\hat\theta_T\right) \right) \right) \right)$$

$$= \frac{1}{T}\frac{1}{H^2} \sum_{h=1}^H \sum_{h'=1}^H \sum_{t=13}^T \sum_{s=13}^T \mathrm{E}^* \left(\mathbf{Y}_t^{*\Delta}\left(\hat\theta_T\right) \left(R_{t,h}^{*\Delta}(\hat\theta_T) - \mathbf{Y}_t^{*\mathsf{T}\Delta}\left(\hat\theta_T\right) \hat\vartheta_{1,T,h}^\Delta\left(\hat\theta_T\right) \right) \right)$$

$$= \frac{1}{T}\frac{1}{H^2} \sum_{h=1}^H \sum_{h'=1}^H \sum_{j=13-l}^l \sum_{s=13+l}^{T-l} \widehat{\epsilon}_{t,h}^\Delta \widehat{\epsilon}_{t+j,h}^\Delta \mathbf{Y}_t^\Delta\left(\hat\theta_T\right) \mathbf{Y}_{t+j}^\Delta\left(\hat\theta_T\right) + o_p(1)$$

$$= \text{Avar}\left(\frac{1}{H} \sum_{h=1}^H \left(\frac{1}{\sqrt{T}} \sum_{t=13}^T \mathbf{Y}_t^\Delta(\theta_0)\left(R_{t,h}^\Delta(\theta_0) - \mathbf{Y}_t^{\mathsf{T}\Delta}(\theta_0)\vartheta_{1,0}(\theta_0) \right) \right) \right) + o_p(1),$$

where $\widehat{\epsilon}_{t,h}^\Delta = \mathbf{Y}_t^\Delta\left(\hat\theta_T\right)\left(R_{t,h}^\Delta(\hat\theta_T) - \mathbf{Y}_t^{\mathsf{T}\Delta}\left(\hat\theta_T\right) \hat\vartheta_{1,T,h}^\Delta\left(\hat\theta_T\right) \right)$. This proves (40). Finally, under the parameters restrictions of Appendix A

$$\mathrm{E}^* \left(\sqrt{T} \left(\frac{1}{H} \sum_{h=1}^H \left(\hat\vartheta_{1,T,h}^{*\Delta}\left(\hat\theta_T\right) - \hat\vartheta_{1,T,h}^\Delta\left(\hat\theta_T\right) \right) \right)^{2+\delta} \right)$$

$$= \left(\frac{1}{H} \sum_{h=1}^H \left(\frac{1}{\sqrt{T}} \sum_{t=13}^T \mathbf{Y}_t^\Delta\left(\hat\theta_T\right) \left(R_{t,h}^\Delta(\hat\theta_T) - \mathbf{Y}_t^{\mathsf{T}\Delta}\left(\hat\theta_T\right) \hat\vartheta_{1,T,h}^\Delta\left(\hat\theta_T\right) \right) \right) \right)^{2+\delta} = O_p(1)$$

Proof of Proposition 7: By the first order conditions and a mean value expansion around $\hat\lambda_\Upsilon$,

$$0 = \nabla_\lambda \left(\frac{1}{H} \sum_{h=1}^H \hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \hat\lambda_\Upsilon^* \right) \right)^\mathsf{T} \left(\frac{1}{H} \sum_{h=1}^H \left(\hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \hat\lambda_\Upsilon^* \right) - \hat\psi_{\Upsilon,h}\left(\hat\phi_T, \hat\theta_T, \hat\lambda_\Upsilon \right) \right) - \left(\tilde{\psi}_\Upsilon^* - \tilde\psi_\Upsilon \right) \right)$$

$$= \nabla_\lambda \left(\frac{1}{H} \sum_{h=1}^H \hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \hat\lambda_\Upsilon^* \right) \right)^\mathsf{T} \left(\frac{1}{H} \sum_{h=1}^H \left(\hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \hat\lambda_\Upsilon \right) - \hat\psi_{\Upsilon,h}\left(\hat\phi_T, \hat\theta_T, \hat\lambda_\Upsilon \right) \right) - \left(\tilde{\psi}_\Upsilon^* - \tilde\psi_\Upsilon \right) \right)$$

$$+ \left(\nabla_\lambda \left(\frac{1}{H} \sum_{h=1}^H \hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \hat\lambda_\Upsilon^* \right) \right)^\mathsf{T} \nabla_\lambda \left(\frac{1}{H} \sum_{h=1}^H \hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \bar\lambda_\Upsilon^* \right) \right) \right) \left(\hat\lambda_\Upsilon^* - \hat\lambda_\Upsilon \right),$$

where $\bar\lambda_\Upsilon^* \in \left(\hat\lambda_\Upsilon^*, \hat\lambda_\Upsilon \right)$. Thus,

$$\sqrt{\Upsilon}\left(\hat\lambda_\Upsilon^* - \hat\lambda_\Upsilon \right)$$

$$= - \left(\nabla_\lambda \left(\frac{1}{H} \sum_{h=1}^H \hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \hat\lambda_\Upsilon^* \right) \right)^\mathsf{T} \nabla_\lambda \left(\frac{1}{H} \sum_{h=1}^H \hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \bar\lambda_\Upsilon^* \right) \right) \right)^{-1}$$

$$\times \nabla_\lambda \left(\frac{1}{H} \sum_{h=1}^H \hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \hat\lambda_\Upsilon^* \right) \right)^\mathsf{T} \sqrt{\Upsilon} \left(\frac{1}{H} \sum_{h=1}^H \left(\hat\psi_{\Upsilon,h}^* \left(\hat\phi_T^*, \hat\theta_T^*, \hat\lambda_\Upsilon \right) - \hat\psi_{\Upsilon,h}\left(\hat\phi_T, \hat\theta_T, \hat\lambda_\Upsilon \right) \right) - \left(\tilde{\psi}_\Upsilon^* - \tilde\psi_\Upsilon \right) \right).$$

From the proof of Proposition 5 and Proposition 6, and recalling that $\Upsilon/T \to \pi$, $0 < \pi < 1$,

$$\sqrt{\Upsilon}\left(\hat{\phi}_T^* - \phi_0\right) = \sqrt{\Upsilon}\left(\hat{\phi}_T^* - \hat{\phi}_T\right) + \sqrt{\Upsilon}\left(\hat{\phi}_T - \phi_0\right) = O_p^*(1)$$

$$\sqrt{\Upsilon}\left(\hat{\theta}_T^* - \theta_0\right) = \sqrt{\Upsilon}\left(\hat{\theta}_T^* - \hat{\theta}_T\right) + \sqrt{\Upsilon}\left(\hat{\theta}_T - \theta_0\right) = O_p^*(1).$$

Hence, by the uniform law of large numbers

$$\hat{\lambda}_\Upsilon^* - \lambda_0 = \left(\hat{\lambda}_\Upsilon^* - \hat{\lambda}_\Upsilon\right) + \left(\hat{\lambda}_\Upsilon - \lambda_0\right) = o_p^*(1) + o_p(1) = o_p^*(1)$$

and thus

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

By a similar argument as in the proof of Proposition 4,

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

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

$$+ \frac{1}{H}\sum_{h=1}^H \nabla_\phi \left(\hat{\psi}_{\Upsilon,h}^*\left(\bar{\phi}_T^*, \hat{\theta}_T^*, \hat{\lambda}_\Upsilon\right)\right)^\intercal \sqrt{\Upsilon}\left(\hat{\phi}_T^* - \hat{\phi}_T\right)$$

$$+ \frac{1}{H}\sum_{h=1}^H \nabla_\theta \left(\hat{\psi}_{\Upsilon,h}^*\left(\hat{\phi}_T, \hat{\theta}_T^*, \hat{\lambda}_\Upsilon\right)\right)^\intercal \sqrt{\Upsilon}\left(\hat{\theta}_T^* - \hat{\theta}_T\right) + o_p^*(1)$$

$$= \sqrt{\Upsilon}\left(\frac{1}{H}\sum_{h=1}^H \left(\hat{\psi}_{\Upsilon,h}^*\left(\phi_0, \theta_0, \lambda_0\right) - \hat{\psi}_{\Upsilon,h}\left(\phi_0, \theta_0, \lambda_0\right)\right)\right)$$

$$+ \sqrt{\Upsilon}F_{\phi_0}^\intercal\left(\hat{\phi}_T^* - \hat{\phi}_T\right) + \sqrt{\Upsilon}F_{\theta_0}^\intercal\left(\hat{\theta}_T^* - \hat{\theta}_T\right) + o_p^*(1)$$

Thus, by the same argument used in the proof of Proposition 5 and Proposition 6, we can show that

$$\mathrm{E}^*\left(\sqrt{\Upsilon}\left(\hat{\lambda}_\Upsilon^* - \hat{\lambda}_\Upsilon\right)\right) = o_p(1)$$

and

$$\text{var}^* \left(\sqrt{\Upsilon} \left(\hat{\lambda}^*_\Upsilon - \hat{\lambda}_\Upsilon \right) \right)$$

$$= \text{var}^* \left(\sqrt{\Upsilon} \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\psi}^*_{\Upsilon,h} \left(\phi_0, \theta_0, \lambda_0 \right) - \hat{\psi}_{\Upsilon,h} \left(\phi_0, \theta_0, \lambda_0 \right) \right) \right) \right.$$

$$\left. + \sqrt{\Upsilon} F^{\intercal}_{\phi_0} \left(\hat{\phi}^*_T - \hat{\phi}_T \right) + \sqrt{\Upsilon} F^{\intercal}_{\theta_0} \left(\hat{\theta}^*_T - \hat{\theta}_T \right) - \left(\tilde{\boldsymbol{\psi}}^*_\Upsilon - \tilde{\boldsymbol{\psi}}_\Upsilon \right) \right)$$

$$= \text{Avar} \left(\sqrt{\Upsilon} \left(\frac{1}{H} \sum_{h=1}^{H} \left(\hat{\psi}_{\Upsilon,h} \left(\phi_0, \theta_0, \lambda_0 \right) - \psi_0 \left(\phi_0, \theta_0, \lambda_0 \right) \right) \right) \right.$$

$$\left. + \sqrt{\Upsilon} F^{\intercal}_{\phi_0} \left(\hat{\phi}_T - \phi_0 \right) + \sqrt{\Upsilon} F^{\intercal}_{\theta_0} \left(\hat{\theta}_T - \theta_0 \right) - \left(\tilde{\boldsymbol{\psi}}_\Upsilon - \boldsymbol{\psi} \right) \right) + o_p(1).$$

and by Minkowski inequality, $\text{E}^* \left(\left(\sqrt{\Upsilon} \left(\hat{\lambda}^*_\Upsilon - \hat{\lambda}_\Upsilon \right) \right)^{2+\delta} \right) = O^*_p(1)$, for some $\delta > 0$.

References

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

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.

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

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*.

Barndorff-Nielsen, O.E. and N. Shephard, 2007, "Variation, Jumps, Market Frictions and High Frequency Data in Financial Econometrics," forthcoming in Blundell, R., P. Torsten and W.K. Newey (Eds.): *Advances in Economics and Econometrics. Theory and Applications*, Ninth World Congress, Econometric Society Monographs, Cambridge University Press.

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

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

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

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.

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, 2004, "Variance Risk Premia," Working Paper, Bloomberg L.P. and Courant Institute.

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.

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

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.

Fornari, F. and A. Mele, 2006, "Financial Volatility and Real Economic Activity," working paper ECB and LSE.

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

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.

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.

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.

Tables

Table 1
Parameter estimates for the macroeconomic factors

	Estimate
κ_1	0.0331
μ_1	1.0379
α_1	$2.2206 \cdot 10^{-4}$
β_1	$-9.6197 \cdot 10^{-7}$
κ_2	0.5344
μ_2	1.0415
α_2	0.0540
β_2	-0.0497
$\bar{\kappa}_1$	-0.2992
$\bar{\kappa}_2$	1.2878

Table 2
Parameter estimates for the stock price process and the unobservable factor

	Estimate
s_0	0.1279
s_1	0.0998
s_2	2.5103
s_3	0.2215
κ_3	0.0091
μ_3	0.0493
α_3	2.3023
β_3	0.2055

Table 3
Parameter estimates for the risk-premium process

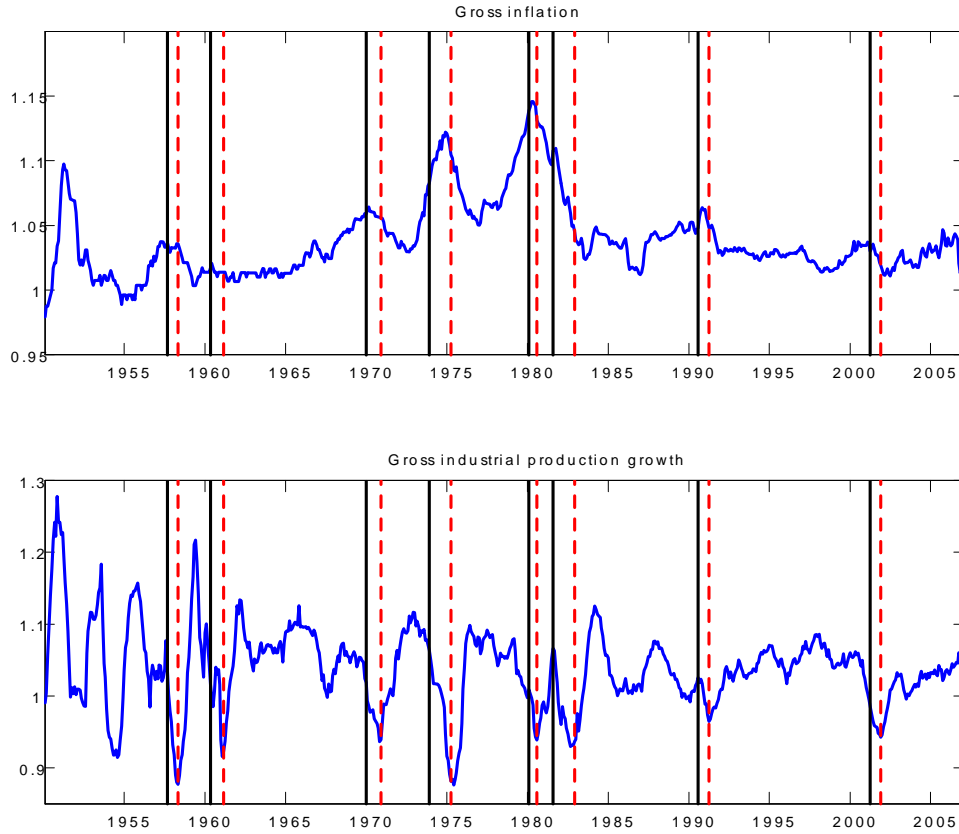|	Estimate ($\div 10^3$)
$\lambda_{1(1)}$	6.4605
$\lambda_{2(1)}$	-0.1159
$\lambda_{1(2)}$	1.4022
$\lambda_{2(2)}$	$2.3332 \cdot 10^{-4}$
$\lambda_{1(3)}$	$-2.2079 \cdot 10^{-6}$
$\lambda_{2(3)}$	$3.7559 \cdot 10^{-6}$

Figures

Figure 1 – Industrial production growth and inflation, with NBER dated recession periods. This figure plots the one-year, monthly gross inflation, defined as $y_{1,t} \equiv \mathrm{CPI}_t / \mathrm{CPI}_{t-12}$, and the one-year, monthly gross industrial production growth, defined as $y_{2,t} \equiv \mathrm{IP}_t / \mathrm{IP}_{t-12}$, where CPI_t is the Consumer price index as of month t, and IP_t is the real, seasonally adjusted industrial production index as of month t. The sample covers monthly data for the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions.
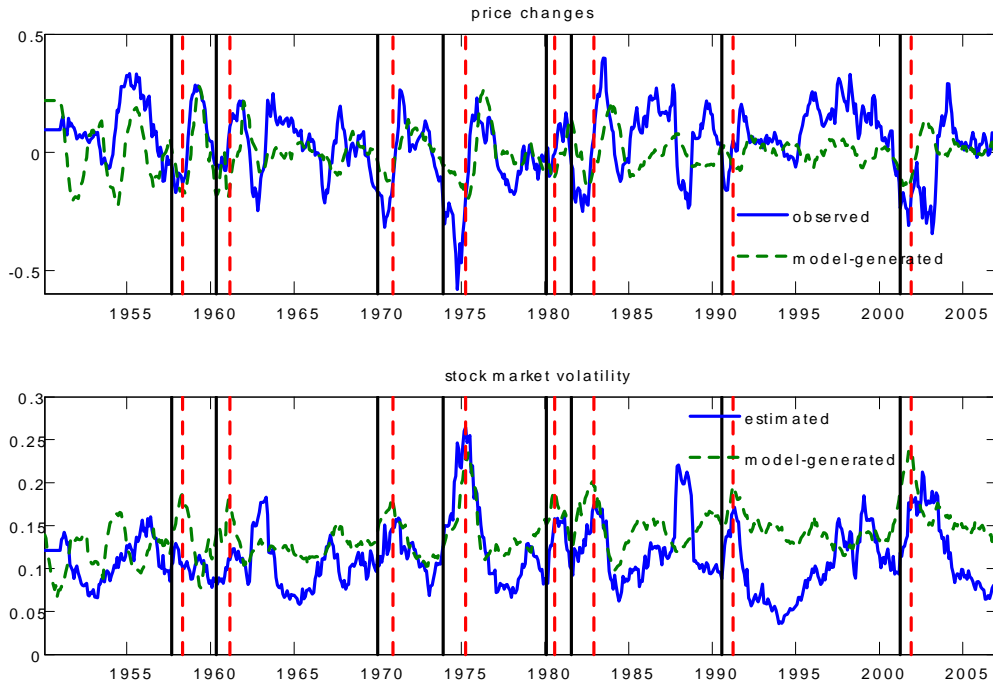
Figure 2 – Returns and volatility along with the model predictions, with NBER dated recession periods. This figure plots one-year ex-post price changes and one-year return volatility, along with their counterparts predicted by the model. The top panel depicts continuously compounded price changes, defined 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. (4). 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.
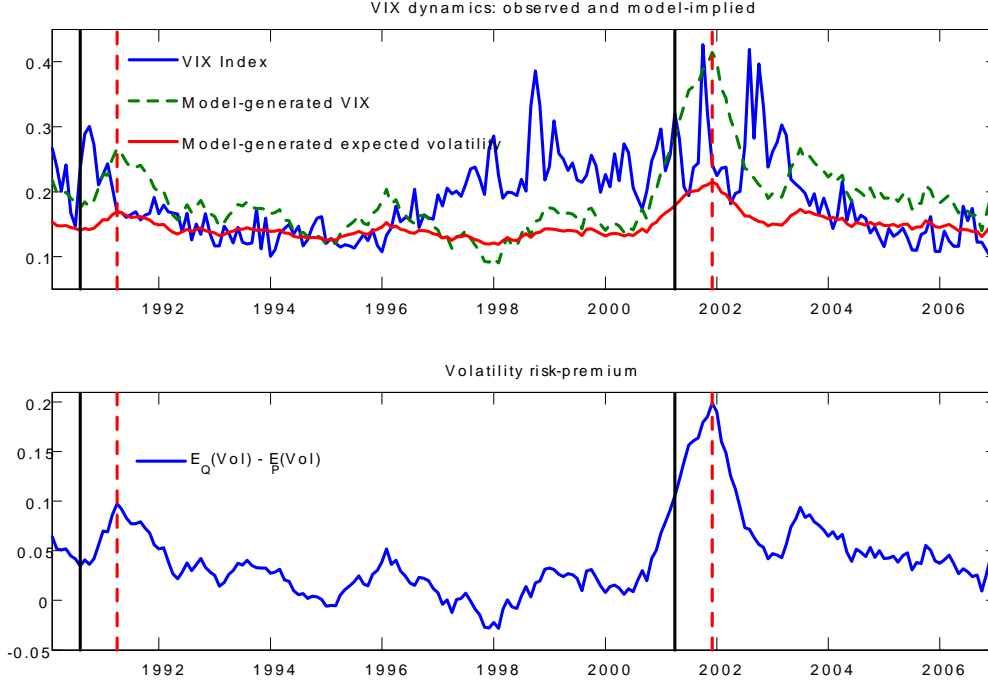
43

Figure 3 – **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,

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

where $T - t = 12^{-1}$, E_Q is the conditional expectation under the risk-neutral probability, $\sigma^2\left(\boldsymbol{y}\right)$ is the instantaneous variance predicted by the model, obtained through Eq. (4), 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.
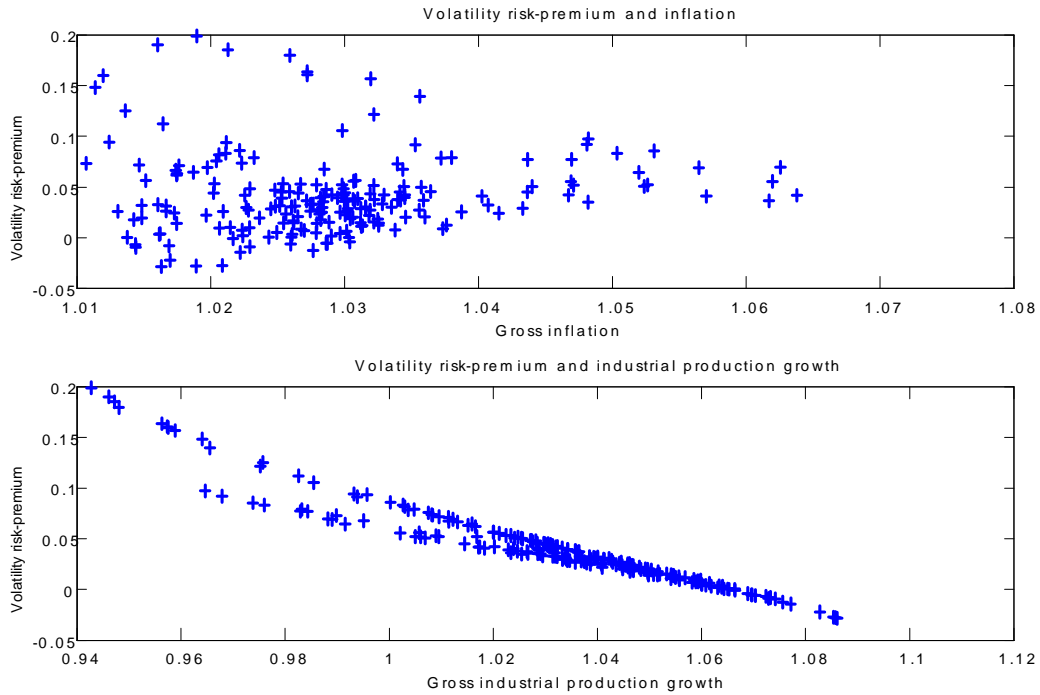
Figure 4 – Volatility risk-premium against inflation and industrial production growth. This figure provides scatterplots of the volatility risk-premium predicted by the model, depicted in Figure 3 (bottom panel), against the two macroeconomic factors depicted in Figure 1 (inflation and growth). Each 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.