Supplemental material for:

Macroeconomic Determinants of Stock Volatility and Volatility Premiums

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

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

[c] *Swiss Finance Institute and University of Lugano*

November 12, 2012.

Remark on equation numbering. The equations in this supplemental material are numbered with (A.1), (A.2), etc. All other equations we refer to here are those in the article.

A. Supplemental material for Section 2

A multifactor model

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

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

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

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

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

for some scalars α_i and vectors $\boldsymbol{\beta}_i$. We assume that the Brownian motion driving secular growth, $W_G(t)$ in Eq. (4), is uncorrelated with $\boldsymbol{W}(t)$ in Eq. (A.1). We shall review soon sufficient conditions known to ensure that Eq. (A.1) has a strong solution with $\boldsymbol{V}(\boldsymbol{y}(t))_{(ii)} > 0$ almost surely for all t.

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

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

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

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

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

where the collection of functions $\{f_j\}$ determines how the factors dynamics impinge upon the observed macroeconomic variables. In terms of the model of the article, the functions in Eq. (A.2) are $f_j(\boldsymbol{y}) \equiv \ln y_j$.

We now turn to model asset prices. We assume that asset prices are related to the vector of factors $\boldsymbol{y}(t)$ in Eq. (A.1), and that some of these factors affect developments in macroeconomic conditions, through Eq. (A.2). For analytical convenience, we rule out that asset prices can feed back the real economy, although we acknowledge that the presence of frictions can make capital markets and the macroeconomy intimately related, as in the financial accelerator hypothesis reviewed by Bernanke, Gertler and Gilchrist (1999), or in the static model analyzed by Angeletos, Lorenzoni and Pavan (2010), where feedbacks arise due to asymmetric information and learning between agents acting within the real and the financial spheres of the economy.

The Arrow-Debreu density we consider is exactly that in Eq. (5), with the sole exception that the vector Brownian motion $\boldsymbol{W}$ is the d-dimensional one in Eq. (A.1). Consider, then, the following "essentially affine" specification for the dynamics of the factors in Eq. (A.1), and the risk-premiums. Let $\boldsymbol{V}^-(\boldsymbol{y})$ be a $d \times d$ diagonal matrix with elements

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

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

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

$$\frac{\mathrm{d}S(t)}{S(t)} = \left(r - \frac{G(t)\,\delta(\boldsymbol{y}(t))}{S(\boldsymbol{y}(t))}\right)\mathrm{d}t + \frac{s_y(\boldsymbol{y}(t))^\top \boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y}(t))}{s(\boldsymbol{y}(t))}\mathrm{d}\hat{\boldsymbol{W}}(t) + \sigma_G \mathrm{d}\hat{W}_G(t), \tag{A.3}$$

where $\hat{\boldsymbol{W}}$ and $\hat{W}_G$ are Brownian motions defined under the risk-neutral probability Q. Under regularity conditions provided below, and in the absence of bubbles, Eq. (A.3) implies that the stock price is,

$$S(G,\boldsymbol{y}) = \mathbb{E}\left[\int_t^\infty e^{-r(s-t)} G(s)\,\delta(\boldsymbol{y}(s))\,\mathrm{d}s \middle| G(t) = G, \boldsymbol{y}(t) = \boldsymbol{y}\right], \tag{A.4}$$

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

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

$$\delta(\boldsymbol{y}) = \delta_0 + \boldsymbol{\delta}^\top \boldsymbol{y}, \tag{A.5}$$

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

We have:

Proposition A1: *Let the risk-premiums be as in Eq. (6), and the instantaneous dividend rate be as in Eqs. (2) and (A.5). Then, under a technical regularity condition (condition (A.10)), we have that: (i) Eq. (A.4) holds; and (ii) the rational stock price function $S(G,\boldsymbol{y}) = G \cdot s(\boldsymbol{y})$, where $s(\boldsymbol{y})$ is affine in the state vector $\boldsymbol{y}$, viz*

$$s(\boldsymbol{y}) = \frac{\delta_0 + \boldsymbol{\delta}^\top \left(\boldsymbol{D} + (r - g + \sigma_G\lambda_G)\boldsymbol{I}_{n\times n}\right)^{-1}\boldsymbol{c}}{r - g + \sigma_G\lambda_G} + \boldsymbol{\delta}^\top \left(\boldsymbol{D} + (r - g + \sigma_G\lambda_G)\boldsymbol{I}_{n\times n}\right)^{-1}\boldsymbol{y}, \tag{A.6}$$

where

$$\boldsymbol{c} = \boldsymbol{\kappa\mu} - \boldsymbol{\Sigma}\begin{pmatrix} \alpha_1\lambda_{1(1)} & \cdots & \alpha_d\lambda_{1(d)} \end{pmatrix}^\top \tag{A.7}$$

$$\boldsymbol{D} = \boldsymbol{\kappa} + \boldsymbol{\Sigma}\left[\begin{pmatrix} \lambda_{1(1)}\boldsymbol{\beta}_1^\top & \cdots & \lambda_{1(d)}\boldsymbol{\beta}_d^\top \end{pmatrix}^\top + \boldsymbol{I}^-\boldsymbol{\lambda}_2\right], \tag{A.8}$$

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

Existence of a strong solution to Eq. (A.1)

Consider the following conditions: for all i,

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

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

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

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

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

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

$$\kappa_i(\mu_i\beta_i + \alpha_i) + \bar{\kappa}_i\beta_i\left(\mu_j + \frac{\alpha_j}{\beta_j}\right) > \frac{1}{2}\beta_i^2, \quad i \neq j. \tag{A.9}$$

Proof of Proposition A1

The technical condition in Proposition A1 is,

$$E\left[\int_t^T \left\| \frac{\boldsymbol{\eta}^\top \boldsymbol{\Sigma} \boldsymbol{V}(\boldsymbol{y}(\tau))}{\gamma + \boldsymbol{\eta}^\top \boldsymbol{y}(\tau)} - \boldsymbol{\Lambda}(\tau)^\top \right\|^2 \mathrm{d}\tau \right] < \infty, \tag{A.10}$$

for some constants γ and $\boldsymbol{\eta}$ in Eq. (A.19) below.

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

$$g = \sigma_G \equiv 0. \tag{A.11}$$

Then, we generalize, by elaborating on Eq. (A.4), as in Eq. (A.20) below.

When Eq. (A.11) holds true, define the Arrow-Debreu adjusted asset price process as, $s^\xi(t) \equiv e^{-rt}\xi(t)\, s(\boldsymbol{y}(t))$, $t > 0$. By Itô's lemma, it satisfies,

$$\frac{\mathrm{d}s^\xi(t)}{s^\xi(t)} = \mathrm{Dr}(\boldsymbol{y}(t))\,\mathrm{d}t + \left(\boldsymbol{Q}(\boldsymbol{y}(t))^\top - \boldsymbol{\Lambda}(\boldsymbol{y}(t))^\top\right) \mathrm{d}\boldsymbol{W}(t), \tag{A.12}$$

where

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

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

$$s^\xi(t) = E\left[\int_t^T \delta^\xi(h)\,\mathrm{d}h \middle| \mathbb{F}(t)\right] + E[s^\xi(T) \mid \mathbb{F}(t)], \tag{A.13}$$

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

$$\lim_{T\to\infty} E[s^\xi(T) \mid \mathbb{F}(t)] = 0, \tag{A.14}$$

from which Eq. (A.4) follows, once we show that $\int_t^\infty E[\delta^\xi(h)]\mathrm{d}h < \infty$.

Next, by Eq. (A.13),

$$0 = \frac{\mathrm{d}}{\mathrm{d}\tau} E[s^\xi(\tau) \mid \mathbb{F}(t)]\bigg|_{\tau=t} + \delta^\xi(t). \tag{A.15}$$

Below, we show that

$$E[s^\xi(T) \mid \mathbb{F}(t)] = s^\xi(t) + \int_t^T \mathrm{Dr}(\boldsymbol{y}(h))\, s^\xi(h)\,\mathrm{d}h. \tag{A.16}$$

Therefore, by the assumptions on $\boldsymbol{\Lambda}$, Eq. (A.15) can be rearranged to yield the following ordinary differential equation,

$$\text{For all } \boldsymbol{y}, \quad s_y(\boldsymbol{y})^\top (\boldsymbol{c} - \boldsymbol{D}\boldsymbol{y}) + \frac{1}{2}\mathrm{Tr}\left([\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})][\boldsymbol{\Sigma}\boldsymbol{V}(\boldsymbol{y})]^\top s_{yy}(\boldsymbol{y})\right) + \delta(\boldsymbol{y}) - rs(\boldsymbol{y}) = 0, \tag{A.17}$$

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

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

$$s(\boldsymbol{y}) = \gamma + \boldsymbol{\eta}^\top \boldsymbol{y}, \tag{A.18}$$

for some scalar γ and some vector $\boldsymbol{\eta}$. By plugging this guess back into Eq. (A.17) we obtain,

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

That is,

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

The solution to this system is,

$$\gamma = \frac{\delta_0 + \boldsymbol{\eta}^\top \boldsymbol{c}}{r} \quad \text{and} \quad \boldsymbol{\eta}^\top = \boldsymbol{\delta}^\top \left(\boldsymbol{D} + r\boldsymbol{I}_{n\times n}\right)^{-1}. \tag{A.19}$$

We are left to show that Eq. (A.14) and (A.16) hold true, when Eq. (A.11) also holds true.

As for Eq. (A.14), we have:

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

where the second line follows by Eq. (A.18), and the third line holds because $E[\xi(T) \mid \mathbb{F}(t)] = 1$ for all T, and by a change of measure. Eq. (A.14) follows because $\boldsymbol{y}$ is stationary mean-reverting under the risk-neutral probability.

To show that Eq. (A.16) holds, we need to show that the diffusion part of s^ξ in Eq. (A.12) is a martingale, not only a local martingale, which it does whenever for all T,

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

which is the condition in (A.10). This ends the proof of Proposition A1, in the case $g = \sigma_G \equiv 0$.

For the general case of Proposition A1, note that by Eq. (A.4):

$$\begin{aligned}
&S(G, \boldsymbol{y}) \\
&= G\cdot \mathbb{E}\left[\int_t^\infty e^{-r(s-t)} \mathbb{E}\left(\frac{G(s)}{G}\delta(\boldsymbol{y}(s))\,\mathrm{d}s \,\middle|\, G(t) = G\right) \middle|\, G(t) = G, \boldsymbol{y}(t) = \boldsymbol{y}\right] \\
&= G\cdot \mathbb{E}\left[\int_t^\infty e^{-r(s-t)} \mathbb{E}\left(e^{\left(g - \frac{1}{2}\sigma_G^2 - \lambda_G\sigma_G\right)(s-t) + \sigma_G\left(\tilde{W}_G(s) - \tilde{W}_G(t)\right)} \,\middle|\, G(t) = G\right)\delta(\boldsymbol{y}(s))\,\mathrm{d}s \,\middle|\, \boldsymbol{y}(t) = \boldsymbol{y}\right] \\
&= G\cdot \mathbb{E}\left[\int_t^\infty e^{-(r - g + \lambda_G\sigma_G)(s-t)}\delta(\boldsymbol{y}(s))\,\mathrm{d}s \,\middle|\, \boldsymbol{y}(t) = \boldsymbol{y}\right],
\end{aligned} \tag{A.20}$$

where the first equality follows by the law of iterated expectations, the second by the independence of G and $\boldsymbol{y}$, and the definition of G in Eq. (4), and the third from a simple computation. The term in the brackets is the same as the RHS of Eq. (A.4), for $G(s) \equiv 1$, $s \in (t, \infty)$. Therefore, the solution for the term in the brackets is the same as that provided in the case of absence of secular growth, i.e. when Eq. (A.11) holds true, but with $r - g + \lambda_G\sigma_G$ replacing r.

B. Supplemental material for Section 3

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

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

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

$$\boldsymbol{\Phi} = \{\boldsymbol{\phi} : \text{The inequality in (A.9) holds, } \kappa_i > 0, \text{ and } \kappa_i\kappa_j - \bar{\kappa}_i\bar{\kappa}_j > 0,\ i,j = 1,2 \text{ and } i \neq j\},$$

and

$$\boldsymbol{\Theta} = \{\boldsymbol{\theta} : \text{The inequality in (A.9) holds for } i = 3, \text{ and } \kappa_3 > 0\}.$$

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

$$\boldsymbol{\phi}_0 = \arg \min_{\boldsymbol{\phi} \in \boldsymbol{\Phi}_0} \operatorname*{plim}_{T \to \infty, \Delta \to 0} \left\| \frac{1}{H} \sum_{h=1}^{H} \hat{\boldsymbol{\varphi}}_{T,\Delta,h} (\boldsymbol{\phi}) - \tilde{\boldsymbol{\varphi}}_T \right\|^2, \tag{A.21}$$

and

$$\boldsymbol{\theta}_0 = \arg \min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}_0} \operatorname*{plim}_{T \to \infty, \Delta \to 0} \left\| \frac{1}{H} \sum_{h=1}^{H} \hat{\boldsymbol{\vartheta}}_{T,\Delta,h} (\boldsymbol{\theta}) - \tilde{\boldsymbol{\vartheta}}_T \right\|^2, \tag{A.22}$$

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

$$\boldsymbol{\lambda}_0 = \arg \min_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_0} \operatorname*{plim}_{T \to \infty, \Delta \to 0} \left\| \frac{1}{H} \sum_{h=1}^{H} \hat{\boldsymbol{\psi}}_{T,\Delta,h}(\hat{\boldsymbol{\phi}}_T, \hat{\boldsymbol{\theta}}_T, \boldsymbol{\lambda}) - \tilde{\boldsymbol{\psi}}_T \right\|^2. \tag{A.23}$$

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

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

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

We have:

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

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

where $\boldsymbol{\phi}_0$ is as in Eq. (A.21), and the two matrices, $\boldsymbol{D}_1$ and $\boldsymbol{J}_1$, are defined in the proof below.

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned} &\sqrt{T}\left(\hat{\boldsymbol{\varphi}}_{1,T,\Delta,h}(\boldsymbol{\phi}_0) - \hat{\boldsymbol{\varphi}}_{1,T,h}(\boldsymbol{\phi}_0)\right) \\ &= \left(\frac{1}{T} \sum_{t=25}^T \boldsymbol{Y}_{t,h}^{\phi_0} \boldsymbol{Y}_{t,h}^{\phi_0\top} \right)^{-1} \sqrt{T} \left(\frac{1}{T} \sum_{t=25}^T \left(\boldsymbol{Y}_{t,\Delta,h}^{\phi_0} y_{1,t,\Delta,h}^{\phi_0} - \boldsymbol{Y}_{t,h}^{\phi_0} y_{1,t,h}^{\phi_0} \right) \right) \\ &+ \sqrt{T} \left(\left(\frac{1}{T} \sum_{t=25}^T \boldsymbol{Y}_{t,\Delta,h}^{\phi_0} \boldsymbol{Y}_{t,\Delta,h}^{\phi_0\top} \right)^{-1} - \left(\frac{1}{T} \sum_{t=25}^T \boldsymbol{Y}_{t,h}^{\phi_0} \boldsymbol{Y}_{t,h}^{\phi_0\top} \right)^{-1} \right) \left(\frac{1}{T} \sum_{t=25}^T \boldsymbol{Y}_{t,\Delta,h}^{\phi_0} y_{1,t,\Delta,h}^{\phi_0} \right). \end{aligned} \tag{A.24}$$

As for the first term on the RHS of (A.24), $\left(\tfrac{1}{T} \sum_{t=25}^T \boldsymbol{Y}_{t,h}^{\phi_0} \boldsymbol{Y}_{t,h}^{\phi_0\top} \right)^{-1} = O_p(1)$, and by Theorem 2.3 in Pardoux and Talay (1985), we have, for $\varepsilon > 0$ and $\sqrt{T}\Delta \to 0$,

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

The second term on the RHS of Eq. (A.24) can be dealt with similarly. Thus, we have:

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

where,

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

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

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

Finally, given Assumption B1(ii),

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

and so

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

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

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

We have:

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

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

where $\boldsymbol{\theta}_0$ *is as in Eq. (A.22), and the four matrices,* $\boldsymbol{D}_2$, $\boldsymbol{J}_2$, $\boldsymbol{K}_2$ *and* $\boldsymbol{P}_2$, *are defined in the proof below.*

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

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

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

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

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

Therefore:

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

where

$$\begin{aligned}\boldsymbol{J}_0 = &\operatorname{Avar}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right)\right) + \operatorname{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right)\\ &+ \boldsymbol{C}_2^\top \operatorname{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T} - \boldsymbol{\theta}_G\right)\right)\boldsymbol{C}_2 + 2\boldsymbol{C}_2^\top \operatorname{Acov}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right), \sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T} - \boldsymbol{\theta}_G\right)\right)\\ &- 2\operatorname{Acov}\left(\sqrt{T}\left(\tfrac{1}{H}\textstyle\sum_{h=1}^H \hat{\boldsymbol{\vartheta}}_{T,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\theta}_G\right) - \boldsymbol{\vartheta}_0\right), \sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right)\\ &- 2\boldsymbol{C}_2^\top \operatorname{Acov}\left(\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{G,T} - \boldsymbol{\theta}_G\right), \sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right).\end{aligned} \tag{A.25}$$

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

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

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

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

where

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

and

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

Therefore, using the fact that $\text{Avar}\left(\sqrt{T}\left(\tilde{\boldsymbol{\vartheta}}_T - \boldsymbol{\vartheta}_0\right)\right) = \text{Avar}\left(\sqrt{T}\left(\hat{\boldsymbol{\vartheta}}_{T,1}\left(\boldsymbol{\theta}_0\right) - \boldsymbol{\vartheta}_0\right)\right) = \boldsymbol{J}_2$, letting $\boldsymbol{P}_2$ denoting the sum of the third, fourth and sixth terms in Eq. (A.25), and exploiting the expression for $\boldsymbol{J}_0$, we obtain:

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

and, hence:

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

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

We construct a simulated series of length $\mathcal{T}$ for the VIX index, at a monthly frequency. Since we do not have a closed-form formula for the VIX index, we need to resort to numerical methods aiming to approximate it. We address this issue by simulating the three factors at a daily frequency, which we then use to numerically integrate the daily volatilities. For each simulation draw $h = 1, \cdots, H$, we initialize each monthly path at the values taken by the observable macroeconomic factors, i.e. at $y_{1,t}$, $y_{2,t}$, $t = T_*, \cdots, T_* + \mathcal{T} - 1$, where T_* is the first date where the VIX is available, and at the monthly unconditional mean of the unobservable factor. In the additional experiments of Appendix C, we initialize each monthly path of the unobservable factor at the values taken by (i) the model-implied factor and (ii) the University of Michigan Consumer Sentiment index to generate the statistics, as defined, respectively, by Eq. (28) and by Eq. (A.36) below. For $i = 1, 2, 3$, $h = 1, \cdots, H$, $k = 0, \cdots, \hat{\Delta}^{-1} - 1$, let $\hat{y}^{\lambda}_{i,t+k\hat{\Delta},h}$ be the value of the i-th factor, at time $t + k\hat{\Delta}$, for the h-th simulation under the risk-neutral probability, performed with parameter $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}_0$ and remaining parameters fixed at their estimates obtained in the first and second step of our estimation procedure. $\hat{\Delta}$ will be defined in a moment. Simulations are obtained through a Milstein approximation to the risk-neutral version of Eq. (1),

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

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

$$\sigma^2_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}) = \hat{\sigma}^2_G + \frac{\sum_{i=1}^{3}\hat{s}^2_{i,T}\left(\hat{\alpha}_{i,T} + \hat{\beta}_{i,T}\hat{y}^{\lambda}_{i,t+k\hat{\Delta},h}\right)}{\tilde{s}^2_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda})}, \tag{A.26}$$

where

$$\tilde{s}_{t+k\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}) = \hat{s}_{0,T} + \sum_{i=1}^{3}\hat{s}_{i,T}\hat{y}^{\lambda}_{i,t+k\hat{\Delta},h}, \tag{A.27}$$

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

$$\text{VIX}_{t,\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda}) = \sqrt{\frac{1}{\hat{\Delta}}\sum_{k=0}^{\hat{\Delta}^{-1}-1}\sigma^2_{t+(k+1)\hat{\Delta},h}(\hat{\boldsymbol{\theta}}_T, \hat{\boldsymbol{\phi}}_T, \hat{\sigma}_{G,T}, \boldsymbol{\lambda})}. \tag{A.28}$$

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

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

We have:

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

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

where $\boldsymbol{\lambda}_0$ *is as in Eq. (A.23), and the four matrices,* $\boldsymbol{D}_3$, $\boldsymbol{J}_3$, $\boldsymbol{K}_3$ *and* $\boldsymbol{P}_3$, *are defined in the proof below.*

Proof: Given Assumptions B1(i) and B1(iii), for any $\boldsymbol{\lambda}$ in a compact set $\boldsymbol{\Lambda}_0$, $y^{\lambda}_{i,t+(k+1)\Delta,h}$, $i = 1,2,3$, $h = 1,\cdots,H$, is geometrically β-mixing, and has a stationary distribution with exponential tails. Thus, by Eqs. (A.26), (A.27) and (A.28), $\mathrm{VIX}_{t,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\phi}_0, \sigma_G, \boldsymbol{\lambda}_0\right)$ is also geometrically β-mixing with exponential tails. Therefore, $\mathrm{VIX}_{t,\Delta,h}\left(\boldsymbol{\theta}_0, \boldsymbol{\phi}_0, \sigma_G, \boldsymbol{\lambda}_0\right)$ has enough finite moments to satisfy sufficient conditions for the law of large numbers and the central limit theorem to apply. Next, note that $\mathrm{VIX}_{t,\Delta,h}\left(\boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_G, \boldsymbol{\lambda}\right)$ is continuously differentiable in the interior of $\boldsymbol{\Phi}_0 \times \boldsymbol{\Theta}_0 \times \boldsymbol{\Sigma}_G \times \boldsymbol{\Lambda}_0$ (for some compact set $\boldsymbol{\Sigma}_G$) and, hence, the uniform law of large numbers also applies. Similarly as in the proof of Propositions B2, we take into account the contribution of parameter estimation error, arising because the risk-neutral paths of the factors are generated using $\hat{\boldsymbol{\phi}}_T$,$\hat{\boldsymbol{\theta}}_T$ and $\hat{\sigma}_{G,T}$, instead of the unknown $\boldsymbol{\phi}_0$, $\boldsymbol{\theta}_0$ and σ_G.

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

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

where

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

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

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

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

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

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

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

where

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

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

and

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

B.2. Bootstrap estimates of the standard errors

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

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

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

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

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

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

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

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

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

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

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

We compute the bootstrap covariance matrix, as follows:

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

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

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

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

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

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

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

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

The Proposition follows, once we show that:

$$\mathrm{E}^*\left(\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T^* - \tilde{\boldsymbol{\varphi}}_T\right)\right) = o_p(1), \tag{A.29}$$

$$\mathrm{var}^*\left(\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T^* - \tilde{\boldsymbol{\varphi}}_T\right)\right) = \mathrm{var}\left(\sqrt{T}\left(\tilde{\boldsymbol{\varphi}}_T - \boldsymbol{\varphi}_0\right)\right) + O_p(l/\sqrt{T}), \tag{A.30}$$

and for $\varepsilon > 0$,

$$\mathrm{E}^*\left(\left(\sqrt{T}\left\|\tilde{\boldsymbol{\varphi}}_T^* - \tilde{\boldsymbol{\varphi}}_T\right\|\right)^{2+\varepsilon}\right) = O_p(1). \tag{A.31}$$

Indeed, under conditions (A.29)-(A.30), we have that by the uniform law of large numbers, $\left|\nabla_\phi\left(\frac{1}{H}\sum_{h=1}^H \hat{\boldsymbol{\varphi}}_{T,h}(\hat{\boldsymbol{\phi}}_T^*)\right) - \boldsymbol{D}_1\right| = o_p^*(1)$. Hence,

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

and, given (A.30), and recalling that $l/\sqrt{T} \to 0$,

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

Given (A.31), the statement follows by Theorem 1 in Goncalves and White (2005).

Let us show (A.29), (A.30) and (A.31). We have,

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

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

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

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

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

This proves (A.29). Next,

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

where $\tilde{\epsilon}_{1,t} = y_{1,t} - \boldsymbol{Y}_t^{\top}\tilde{\boldsymbol{\varphi}}_{1,T}$. This proves (A.30). Finally, as $\frac{1}{T}\sum_{t=25}^{T}\boldsymbol{Y}_t\boldsymbol{Y}_t^{\top}$ is full rank, for a generic constant C, and $\varepsilon > 0$,

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

By Lemma 2.1 in Goncalves and White (2005), $\mathrm{E}\left(\mathrm{E}^*\left\|\frac{1}{\sqrt{T}}\sum_{t=25}^{T}\boldsymbol{Y}_t^{*\top}\left(y_{1,t}^* - \boldsymbol{Y}_t^{*\top}\tilde{\boldsymbol{\varphi}}_{1,T}\right)\right\|^{2+\varepsilon}\right) = O(1)$. Hence, (A.31) follows by Markov inequality.

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

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

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

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

$$\ln S_{t,\Delta,h}^{*\theta}(\hat{\boldsymbol{\theta}}_{G,T}^*) = \ln G_{t,\Delta,h}^{*\hat{\theta}_{G,T}^*} + \ln\left(s_0 + s_1 y_{1,t}^* + s_2 y_{2,t}^* + Z_{t,\Delta,h}^{\theta_u,*}\right), \tag{A.32}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

and:

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

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

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

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

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

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

As mentioned in the article, the model free VIX index series is available only from 1990 and so in the third step we have a sample of length $\mathcal{T}$, instead of length T. Thus, we need to resample $y_{1,t}, y_{2,t}, S_t$ and VIX_t from the shorter sample, using blocksize l and number of blocks B, so that $lB = \mathcal{T}$. Also, we need to resample the unobservable factor from a sample of length $\mathcal{T}$, at the parameter estimate of $\boldsymbol{\theta}_u$ obtained in the previous step, $\hat{y}^{\theta_u *}_{3,t,\Delta,h}$ say. Let $\text{VIX}^*_{t,\Delta,h}(\boldsymbol{y}^*_t; \hat{\boldsymbol{\phi}}^*_T, \hat{\boldsymbol{\theta}}^*_T, \hat{\sigma}^*_{G,T}, \boldsymbol{\lambda})$ be the resampled model-based VIX, according to Eq. (A.27). Finally, let

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

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

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

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

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

Construct the bootstrap covariance matrix, as

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

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

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

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

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

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

C. Supplemental material for Section 4

We provide additional results aiming to: (i) ascertain how the model-implied statistics match those of the data (in Section C.1), and (ii) assess some implications of the model estimates for the dynamics of dividends (in Section C.2). Finally, (iii) we conduct experiments to check whether our variance decompositions and out-of-sample results are robust to the modification of the methodology we use to integrate out the unobserved factor (in Section C.3).

C.1. Model-implied dividends dynamics

We explore the implications our three-step estimation procedure has on the dynamics of the dividends. Consider the system of equations $s_i = f_i(\boldsymbol{s}_{\neq i}, \boldsymbol{\lambda}, \boldsymbol{\phi}, \boldsymbol{\theta}_G, \boldsymbol{\theta}; \boldsymbol{\delta}, \lambda_G)$, $i = 0, 1, 2, 3$, for four functions $f_i(\boldsymbol{s}_{\neq i}, \boldsymbol{\lambda}, \boldsymbol{\phi}, \boldsymbol{\theta}_G, \boldsymbol{\theta}; \boldsymbol{\delta}, \lambda_G)$ given by Eqs. (9)-(10)-(11), where $\boldsymbol{s}_{\neq i}$ denotes the parameter vector that includes all the elements s_j except s_i, and the remaining notation is as in the article. Given the estimates of $\boldsymbol{s}, \boldsymbol{\lambda}, \boldsymbol{\phi}, \boldsymbol{\theta}_G$, and $\boldsymbol{\theta}$, say $\hat{\boldsymbol{s}}$, $\hat{\boldsymbol{\lambda}}$, $\hat{\boldsymbol{\phi}}$, $\hat{\boldsymbol{\theta}}_G$ and

$\hat{\boldsymbol{\theta}}$ reported in Section 4 (see Tables 1, 2 and 4), we set $r = 0.01$, and search for values of $\boldsymbol{\delta}$ and λ_G that jointly minimize the criterion,

$$\Xi(\boldsymbol{\delta}, \lambda_G) \equiv \sum_{i=0}^{3} (\hat{s}_i - f_i(\hat{\boldsymbol{s}}_{\neq i}, \hat{\boldsymbol{\lambda}}, \hat{\phi}, \hat{\boldsymbol{\theta}}_G, \hat{\boldsymbol{\theta}}; \boldsymbol{\delta}, \lambda_G))^2, \tag{A.33}$$

as well as moment conditions relying on an auxiliary model for the real dividends (see Eq. (A.35) below), along with the mean and the standard deviation of real dividends, say $\bar{\delta}$ and $\hat{\sigma}_\delta$.

To generate moment conditions for the dividends, we simulate H paths of length T of the unobservable factor $y_3(t)$, and the unobservable secular growth, $G(t)$, using a Milstein approximation with discrete interval Δ, using the parameter estimates of $\boldsymbol{\theta}_u$ and $\boldsymbol{\theta}_G$ in Table 2, and sample them at the same frequency as the data, obtaining the series $y_{3,t,\Delta,h}^{\theta_u}$ and $G_{t,\Delta,h}^{\theta_G}$. Likewise, let $\text{Div}_{t,\Delta,h}^{\theta_u,\theta_G,\delta}$ be the simulated series of the dividends, when the parameters are fixed at $\boldsymbol{\theta}$:

$$\ln \text{Div}_{t,\Delta,h}^{\theta_u,\theta_G,\delta} = \ln G_{t,\Delta,h}^{\theta_G} + \ln\left(\delta_0 + \delta_1 y_{1,t} + \delta_2 y_{2,t} + \delta_3 y_{3,t,\Delta,h}^{\theta_u}\right), \tag{A.34}$$

where $y_{1,t}$ and $y_{2,t}$ denote gross inflation and gross industrial production growth. Eq. (A.34) is, naturally, the simulated counterpart to Eqs. (2)-(3), with $G_{0,\Delta,h}^{\theta_G} \equiv 1$, as in Eq. (4). We fix the intercept, δ_0, so as to make the model-implied average of the detrended dividends match its empirical counterpart: $\delta_0 = \overline{\text{Div}}^d - \delta_1\bar{y}_1 - \delta_2\bar{y}_2 - \delta_3$, where $\overline{\text{Div}}^d$ denotes the sample mean of the detrended dividends, $\text{Div}_t^d \equiv e^{-gt}\text{Div}_t$, Div_t is the observed real dividend at time t, and finally, $\bar{y}_1$ and $\bar{y}_2$ are the sample means of two macroeconomic factors gross inflation and gross industrial production growth, $y_{1,t}$ and $y_{2,t}$, depicted in Figure 1. Real dividends are defined as dividends divided by the consumer price index, and dividend data are obtained from Robert Shiller's website (`http://www.econ.yale.edu/~shiller/`), covering monthly data for the period from January 1950 to December 2006. Next, define yearly dividend growth as $\text{DG}_t \equiv \ln(\text{Div}_t/\text{Div}_{t-12})$, and let $\text{DG}_{t,\Delta,h}^{\theta_u,\theta_G,\delta}$ be the simulated counterparts to DG_t. The auxiliary model for dividend growth is:

$$\text{DG}_t = a^{\text{D}} + \sum_{i\in\{12,24\}} b_{1,i}^{\text{D}} y_{1,t-i} + \sum_{i\in\{12,24\}} b_{2,i}^{\text{D}} y_{2,t-i} + \epsilon_t^{\text{D}}, \tag{A.35}$$

and is, naturally, the same as that we use to fit the model-implied dividend growth, $\text{DG}_{t,\Delta,h}^{\theta_u,\theta_G,\delta}$. Our optimization leads us to find that $\lambda_G = 139.01$ and the values of $\boldsymbol{\delta}$ reported in Table C.1 below.

Table C.1

Calibrated values of $\boldsymbol{\delta}$, the parameter vector relating to the dividend process in Eqs. (2)-(3), as implied by the three-step estimation procedure.

δ_0	δ_1	δ_2	δ_3
0.0382	−0.0302	0.0291	0.0006

Table C.2 (Panel E) reports parameter estimates for the auxiliary model for the real dividends in Eq. (A.35). Figure C.1 depicts the dynamics of real dividend growth, DG_t, as well as its simulated counterparts, obtained by feeding Eq. (A.34) with the realization of the two macroeconomic factors, gross inflation, $y_{1,t}$, and gross industrial production growth, $y_{2,t}$, and by averaging over the cross-section of 1000 simulations of secular growth and the unobserved factor and, finally, by fixing the parameters to $\hat{\boldsymbol{\theta}}_u$ and $\hat{\boldsymbol{\theta}}_G$ (as reported in Table 2 of Section 4), and the calibrated values for $\boldsymbol{\delta}$ reported in Table C.1.

The calibrated parameter values in Table C.1 suggest real dividend growth is procyclical in our model, at least because it positively links to gross industrial production growth, through the positive coefficient value, δ_2. While it also negatively relates to inflation (due to $\delta_1 < 0$), it is overall procyclical, in that our model-implied dividend growth drops over *all* NBER recession episodes of our sample, mimicking the behavior of their observed counterparts, both qualitatively (especially over the last ten years in the sample, and quantitatively (as seen from the parameter estimates for the auxiliary model in Panel E of Table C.2). Note that the negative coefficient estimates of (the sum of) $b_{2,12}^{\text{D}}$ and $b_{2,24}^{\text{D}}$ are consistent with a procyclical dividends behavior, given a mean-reverting behavior of the industrial production growth. Intuitively, bad times (when industrial production is low) are followed

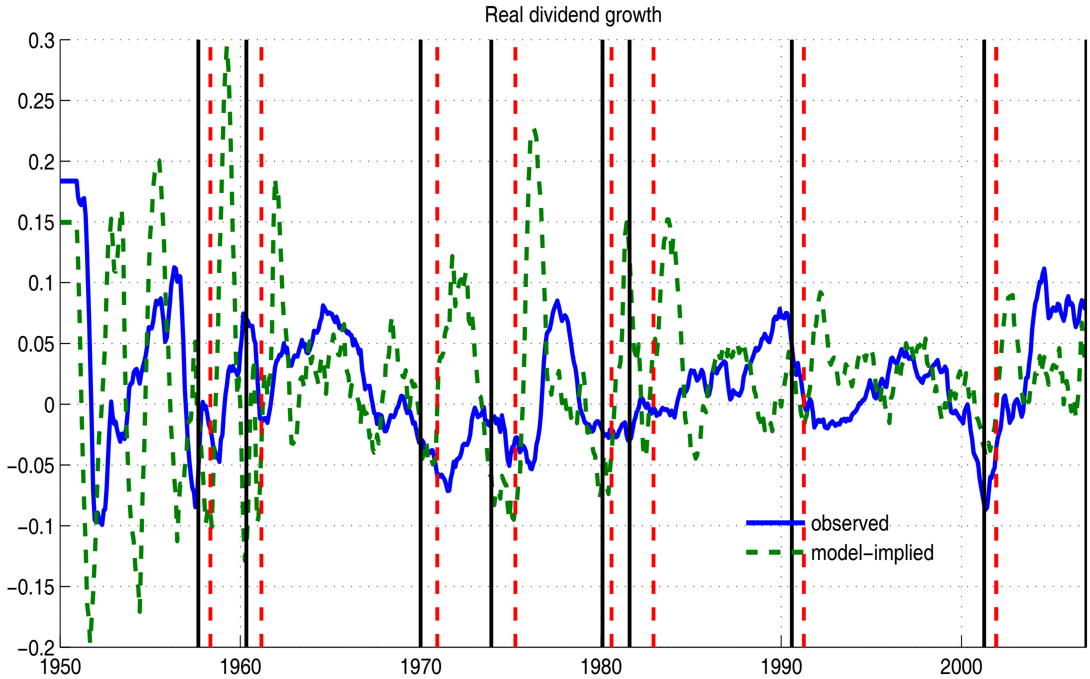

Figure C.1 – Dividend growth, observed and implied by the model estimates, with NBER dated recession periods. This figure plots year-to-year dividend growth, along with its counterpart implied by the model estimates, obtained through our three-step procedure. Dividend growth is defined as $\mathrm{DG}_t = \ln(\mathrm{Div}_t/\mathrm{Div}_{t-12})$, where Div_t is the dividend observed at time t. The model-implied dividend growth is obtained by feeding Eq. (A.34) with the two macroeconomic factors depicted in Figure 1 (inflation and growth), by averaging over the cross-section of 1000 simulations of secular growth and the unobserved factor and, finally, by fixing the parameters to $\hat{\boldsymbol{\theta}}_u$ and $\hat{\boldsymbol{\theta}}_G$ (as reported in Table 2 of Section 4), and the calibrated values for $\boldsymbol{\delta}$ reported in Table C.1. The sample covers monthly data for the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions.

by good. But good times are those where detrended dividends also increase. Therefore, a slowdown in industrial production growth is a predictor of high dividend growth. (A similar explanation holds for the negative values of the macroeconomic loadings resulting from the estimates of the auxiliary regressions relating to the asset returns in Panel B, and the positive values of the short-term macroeconomic loadings relating to the auxiliary regressions for volatility reported in Panel C.)

To illustrate with a simple example, consider a case where detrended dividends positively link to a single state variable tracking the business cycle conditions $x(t)$, say, such that $\delta(t) = \delta_0 + \delta_x x(t)$, for two constant δ_0 and δ_x, and where $\delta_x > 0$. Assume, then, and critically, that $x(t)$ is mean-reverting, with unconditional expectation μ, speed of adjustment $\kappa > 0$, and some volatility coefficient $\sigma(x)$,

$$dx(t) = \kappa(\mu - x(t))\,dt + \sigma(x(t))\,dW(t),$$

where $W(t)$ is a Brownian motion. Then, it is straightforward to show that $E_{t-12}(\delta(t) - \delta(t-12)) = a_0 - a_1 x(t-12)$, where E_t denotes the expectation taken conditionally upon the information set as of time t, and $a_0 \equiv \delta_x \left(1 - e^{-12\kappa}\right) \mu$ and $a_1 \equiv \delta_x \left(1 - e^{-12\kappa}\right)$. That is, if $x(t)$ is mean-reverting, $\kappa > 0$, and $\delta(t)$ is procyclical, $\delta_x > 0$, expected changes in detrended dividends negatively link to past values of $x(t)$, i.e. $a_1 > 0$. This reasoning generalizes to a multivariate case, although the presence of feedbacks between macroeconomic variables might then dilute the contribution of each variable as a predictor of future detrended dividends.

C.2. Model-implied predictions of reduced-form regressions

This section contains details concerning parameter estimates for the auxiliary models utilized to implement the three-step estimation procedure in the main paper—Table C.2, Panel A through D. The values of $\boldsymbol{\delta}$ in Table C.1, which simultaneously minimize the criterion function $\Xi(\boldsymbol{\delta}, \lambda_G)$ in Eq. (A.33) and moment conditions for the auxiliary models for the real dividends in Eq. (A.35), lead to the parameter estimates for the auxiliary models reported in Panel E of Table C.2.

Table C.2

Parameter estimates for the auxiliary models fitted on both data and data generated by the model, and relating to (i) the macroeconomic factors (Eqs. (17)-(18)), in Panel A; (ii) asset returns (Eq. (22)), in Panel B; (iii) asset volatility (Eq. (23)), in Panel C; and the VIX index (Eq. (25)), in Panel D. Panel E reports parameter estimates for the auxiliary model relating to the real dividends (Eq. (A.35)). For each of these auxiliary models, we report R^2 as well as the residual variances, denoted with c_1 and c_2 (Panel A), $\sigma^2_{\epsilon^{\mathrm{R}}}$ (Panel B), $\sigma^2_{\epsilon^{\mathrm{V}}}$ (Panel C), $\sigma^2_{\epsilon^{\mathrm{VIX}}}$ (Panel D), and $\sigma^2_{\epsilon^{\mathrm{D}}}$ (Panel E). The parameter estimates w_i, a_{ij} and c_i in Panel A refer to the vector $\boldsymbol{w}$, the matrix $\boldsymbol{A}$ and the diagonal elements of the variance-covariance matrix $\boldsymbol{C}$ in Eqs. (17)-(18), and $R^2_{y_i}$ is the R^2 of the regression for the macroeconomic factor y_i, $i = 1, 2$. Remaining notation is as in the article.

Panel A

Auxiliary regressions relating to macroeconomic factors

	Data	Model-implied
w_1	−0.0197	−0.0232
w_2	0.1640	0.2134
a_{11}	0.9982	0.9970
a_{21}	−0.1063	−0.1255
a_{12}	0.0208	0.0254
a_{22}	0.9483	0.9199
$R^2_{y_1}$	0.9516	0.9203
c_1	$1.41 \cdot 10^{-5}$	$2.03 \cdot 10^{-5}$
$R^2_{y_2}$	0.9394	0.9114
c_2	0.0002	0.0004
$\bar{y}_1$	1.0385	1.0386
$\bar{y}_2$	1.0367	1.0375
$\hat{\sigma}_{y_1}$	0.0296	0.0296
$\hat{\sigma}_{y_2}$	0.0571	0.0572

Panel B

Auxiliary regressions relating to asset returns

	Data	Model-implied
a^{R}	2.4293	3.4159
$b^{\mathrm{R}}_{1,12}$	−1.2560	−1.0303
$b^{\mathrm{R}}_{2,12}$	−1.0609	−2.2459
R^2	0.1707	0.5221
$\sigma^2_{\epsilon^{\mathrm{R}}}$	0.0204	0.0139
$\bar{S}$	2.8843	2.7915

Table C.2—*continued*

Panel C

Auxiliary regressions relating to asset volatility

	Data	Model-implied
a^{V}	−0.3745	0.0628
b_6^{V}	1.0162	0.8525
b_{12}^{V}	−0.6893	−0.2889
b_{18}^{V}	0.5311	0.3741
b_{24}^{V}	−0.3272	−0.1515
b_{36}^{V}	0.0541	0.0269
b_{48}^{V}	−0.0205	0.0167
$b_{1,12}^{\mathrm{V}}$	0.0813	0.2969
$b_{1,24}^{\mathrm{V}}$	−0.1191	−0.1023
$b_{1,36}^{\mathrm{V}}$	0.1831	−0.1114
$b_{1,48}^{\mathrm{V}}$	−0.0035	−0.0649
$b_{2,12}^{\mathrm{V}}$	0.0435	−0.0957
$b_{2,24}^{\mathrm{V}}$	0.0910	0.0517
$b_{2,36}^{\mathrm{V}}$	0.1022	−0.0215
$b_{2,48}^{\mathrm{V}}$	0.0312	0.0036
R^2	0.6184	0.7551
$\sigma^2_{\epsilon^{\mathrm{V}}}$	0.0006	0.0005
$\overline{\mathrm{Vol}}$	0.1150	0.1305
$\hat{\sigma}_{\mathrm{Vol}}$	0.0401	0.0314

Panel D

Auxiliary regressions relating to the VIX index

	Data	Model-implied
a^{VIX}	−0.4316	−0.0037
b^{VIX}	0.5532	0.9788
$b_{1,36}^{\mathrm{VIX}}$	−0.2800	−0.1318
$b_{1,48}^{\mathrm{VIX}}$	−0.0357	0.0989
$b_{2,36}^{\mathrm{VIX}}$	0.3320	0.0518
$b_{2,48}^{\mathrm{VIX}}$	0.4865	−0.0107
R^2	0.6969	0.9512
$\sigma^2_{\epsilon^{\mathrm{VIX}}}$	0.0013	$3.97 \cdot 10^{-5}$
$\overline{\mathrm{VIX}}$	0.1894	0.2326
$\hat{\sigma}_{\mathrm{VIX}}$	0.0636	0.0278

Panel E

Auxiliary regressions relating to detrended dividends

	Data	Model-implied
a^{D}	0.6753	1.6282
$b_{1,12}^{\mathrm{D}}$	−0.5057	−1.0051
$b_{1,24}^{\mathrm{D}}$	−0.0753	0.6234
$b_{2,12}^{\mathrm{D}}$	0.1043	−1.0894
$b_{2,24}^{\mathrm{D}}$	−0.1598	−0.0780
R^2	0.2352	0.5659
$\sigma^2_{\epsilon^{\mathrm{D}}}$	0.0014	0.0026
$\bar{\delta}$	0.0781	0.0879
$\hat{\sigma}_{\delta}$	0.0136	0.0363

C.3. Variance decomposition and out-of-sample statistics

The experiments of this section produce variance decomposition and out-of-sample statistics under two alternative methodologies regarding the treatment of the unobserved factor. In the article, these statistics originate from simulations of the unobservable factor. Under the first methodology of this section, we feed the model through the time series of the model-implied unobservable factor $\hat{y}_3(t)$, calculated as in Eq. (28) of the article. Figure C.2 depicts conditional variance decompositions, calculated through Eq. (27), and obtained by feeding the model with the two macroeconomic factors, inflation and growth in Figure 1, and replacing the unobserved factor $y_3(t)$ with $\hat{y}_3(t)$. Table C.3 (Panel A) summarizes results applying to the overall sample as well as selected subsamples. Panel B of Table C.3 reports the statistics corresponding to the VIX index. Sample periods are as indicated in Figure C.2.

Under the second methodology, we use the University of Michigan Consumer Sentiment (UMCSENT) index to generate the statistics. Note that the UMCSENT index has a different scale than that of the model-implied unobserved factor. Therefore, we re-scale the index in a way that it has the same average and standard deviation as the model-implied factor over any given sampling period where we conduct the experiments, as follows. Let Sent_t and $\hat{y}_3(t)$ be the UMCSENT index and, as usual, the model-implied unobserved factor. When calculating statistics through the UMCSENT index, we utilize the time series $-\overline{\text{Sent}}_t$, where we define:

$$\overline{\text{Sent}}_t \equiv -E_T(\hat{y}_3) + \frac{\sigma_T(\hat{y}_3)}{\sigma_T(\text{Sent})}\left(\text{Sent}_t - E_T(\text{Sent})\right), \tag{A.36}$$

with $E_T(x)$ and $\sigma_T(x)$ denoting the average and the standard deviation of a given time series x over a certain sampling period of size T. The rationale behind the term $-E_T(\hat{y}_3)$ is that higher realizations of $\hat{y}_3$ are bad news to the stock market, given the negative sign of the s_3 estimate reported in Table 4, as explained in the article. Therefore, according to Eq. (A.36), sample periods where the extracted factor is on average high correspond to periods where our rescaled index, $\overline{\text{Sent}}_t$, is on average low. The sampling periods we consider are (i) from January 1978 to December 2006, for the variance decompositions relating to stock volatility, where $E_T(\hat{y}_3) = 2.302$, $\sigma_T(\hat{y}_3) = 7.915$, and $\sigma_T(\text{Sent}) = 12.2747$, and (ii) from January 1990 to December 2006, for the decomposition statistics relating to the VIX index, where $E_T(\hat{y}_3) = -2.341$, $\sigma_T(\hat{y}_3) = 7.1305$ and $\sigma_T(\text{Sent}) = 12.2184$. Figure C.3 and Table C.4 contain variance decomposition statistics of these experiments, which are the counterparts to those relating to Figure C.2 and Table C.3.

The findings summarized in Tables C.3-C4 and Figures C.2-C.3 closely match those relying on simulations, and reported in the article (Table 3 and Figure 3). The contributions are quite comparable across all the factors, quantitatively. For example, industrial production growth makes the most important contribution to the overall variation in both realized volatility and the VIX index, with its conditional properties being basically the same as those in Figure 3. The contribution of secular growth is, at times, more important than that of the unobserved factor. These cases arise, for example, during the dotcom bubble of the late 1990s, in the experiments relating to the use of the model-implied factor (Figure C.2), or during the subprime events and the 2007 recession (Figure C.3), in the experiments of the UMCSENT index.

The explanations of these cases link to the fact that the variance of secular growth is constant, such that the contribution of growth, $C_G(t)$ in Eq. (27), is inversely related to volatility, $\sigma^2(t)$. The volatility predicted by the model for the dotcom bubble is quite low, and especially so in the experiments based on the model-implied factor, which explains the findings in Figure C.2. Instead, a higher contribution of growth over the subprime events in Figure C.3, relates to the failure of the model to fully capture the surge in realized volatility experienced over that period, once our model is fed wih the UMCSENT index. This fact is confirmed by out-of-sample experiments, discussed below (see Table C.5, and Figures C.4 and C.5), where our model delivers much better results relating to realized volatility predictions, once we use as an input the model-implied unobserved factor rather than the UMCSENT index, and even controlling for a different window over which to evaluate $E_T(\hat{y}_3)$ and $\sigma_T(\hat{y}_3)$ in Eq. (A.36). (Model predictions based on the UMCSENT index remain, however, better than predictions based on OLS.) At the same time, the model delivers the best out-of-sample predictions in terms of the VIX index, once we utilize the UMCSENT index as an input.

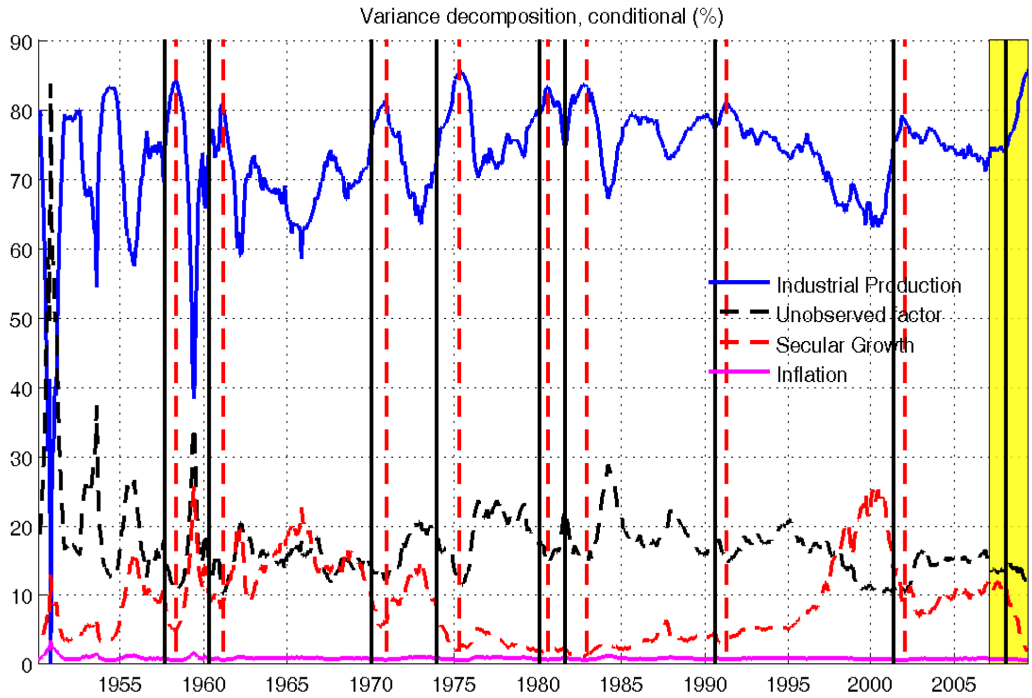

Figure C.2 – Contributions to total stock volatility made by macroeconomic and the model-implied unobservable factors, with NBER dated recession periods. This figure plots the contributions to stock volatility, $C_j(t)$ and $C_G(t)$ in Eq. (27), obtained as the ratios of the instantaneous stock return variance due to factor y_j to the total instantaneous variance, $\sigma^2(t)$, $C_j(t)$ $(j = 1, 2, 3)$, as well as the ratio of the instantaneous variance of secular growth to $\sigma^2(t)$, $C_G(t)$. From top to bottom, "Industrial Production" is $C_2(t)$, "Unobservable factor" is $C_3(t)$, "Secular Growth" is $C_G(t)$, and "Inflation" is $C_1(t)$. Each prediction at each point in time is obtained by feeding the model with the two macroeconomic factors depicted in Figure 1 (inflation and growth), and replacing the unobserved factor with the model-implied factor $\hat{y}_3(t)$, as defined in Eq. (28). The sample covers monthly data for the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions. The shaded area (in yellow) covers the out-of-sample period, from January 2007 to March 2009.

Table C.3

Variance decomposition statistics for (i) realized volatility (Panel A) and expected volatility under the risk-neutral probability (Panel B). Panel A reports averages and standard deviations of the contributions $C_j(t)$ and $C_G(t)$ to the total variance, $\sigma^2(t)$ in Eq. (12), made by: (i) the two macroeconomic factors, gross inflation, $y_1(t)$, and gross industrial production growth, $y_2(t)$, as defined in Table 1, (ii) the model-implied unobserved factor, $\hat{y}_3(t)$, estimated as in Eq. (28),

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

and (iii) secular growth, defined respectively, as:

$$C_j(t)\,\bar{s}^2(\boldsymbol{y}(t)) \equiv \frac{\hat{s}_j^2(\hat{\alpha}_j + \hat{\beta}_j y_j(t))}{\sigma^2(t)}, \; j = 1,2,3, \quad \text{and} \quad C_G(t) \equiv \frac{\sigma_G^2}{\sigma^2(t)},$$

where S_t is the real stock price at time t, $\bar{s}(\boldsymbol{y}) \equiv \hat{s}_0 + \sum_{j=1}^{3} \hat{s}_j y_j$, $\hat{G}_t$ is the cross-sectional average of 1000 simulations of secular growth, $(\hat{\alpha}_j, \hat{\beta}_j)_{j=1}^{3}$ and $(\hat{s}_j)_{j=0}^{3}$ are the parameter estimates, as reported in Tables 1 and 2, $y_3(t) \equiv \hat{y}_{3t}$, and the total variance in Eq. (12) is obtained fixing $y_3(t) \equiv \hat{y}_{3t}$. The sample covers monthly data for the period from January 1950 to December 2006. Panel B reports statistics for the risk-neutral counterparts to the average paths of $C_j(t)$ and $C_G(t)$. The sample covers monthly data for the period from January 1990 to December 2006.

Panel A: Contributions of factors to stock volatility

Averages	1950-2006	1950-1981	1960-1981	1982-2006
Gross inflation	0.87%	0.91%	0.86%	0.82%
Gross growth	73.20%	71.59%	72.63%	74.98%
Unobserved factor	17.35%	18.23%	16.63%	16.39%
Secular growth	8.56%	9.24%	9.85%	7.79%

Standard deviations	1950-2006	1950-1981	1960-1981	1982-2006
Gross inflation	0.22%	0.28%	0.12%	0.09%
Gross growth	7.64%	9.37%	5.87%	4.52%
Unobserved factor	5.94%	7.46%	3.14%	3.36%
Secular growth	5.69%	5.46%	5.71%	5.85%

Panel B: Contributions of factors to the VIX Index

	Averages	Standard deviations
Gross inflation	3.13%	0.03%
Gross growth	85.94%	2.38%
Unobserved factor	8.22%	1.22%
Secular growth	2.71%	0.17%

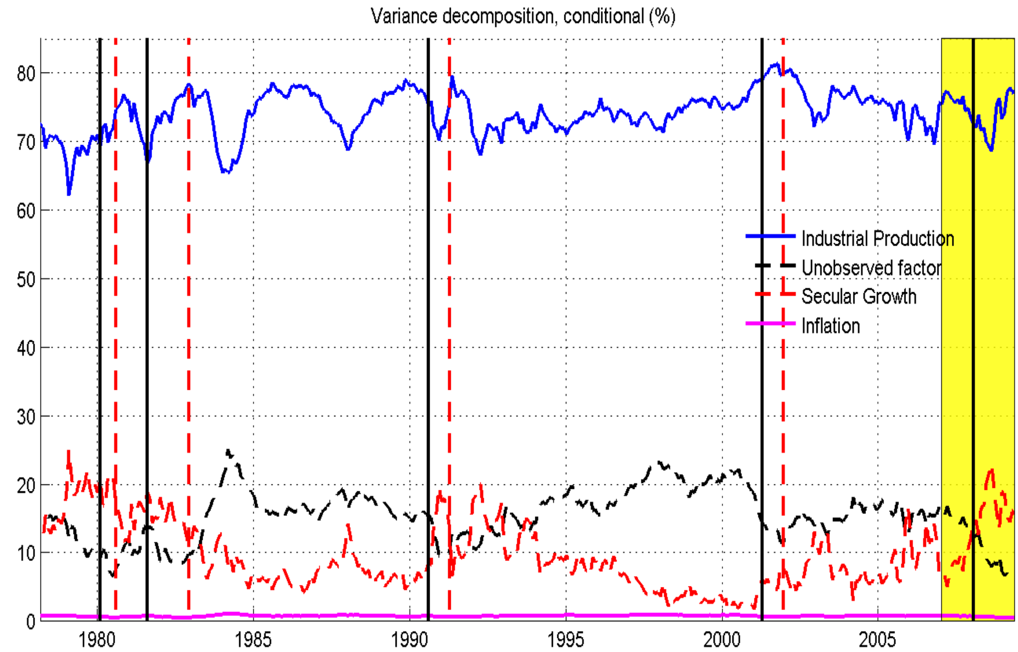

Figure C.3 – Contributions to total stock volatility made by macroeconomic factors and the UMCSENT index, with NBER dated recession periods. This figure plots the contributions to stock volatility, $C_j(t)$ and $C_G(t)$ in Eq. (27), obtained as the ratios of the instantaneous stock return variance due to factor y_j to the total instantaneous variance, $\sigma^2(t)$, $C_j(t)$ $(j = 1, 2, 3)$, as well as the ratio of the instantaneous variance of secular growth to $\sigma^2(t)$, $C_G(t)$. From top to bottom, "Industrial Production" is $C_2(t)$, "Unobservable factor" is $C_3(t)$, "Secular Growth" is $C_G(t)$, and "Inflation" is $C_1(t)$. Each prediction at each point in time is obtained by feeding the model with the two macroeconomic factors depicted in Figure 1 (inflation and growth), and replacing the unobserved factor with the UMCSENT index, rescaled as in Eq. (A.36). The sample covers monthly data for the period from January 1978 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and vertical dashed lines (in red) indicate the end of NBER-dated recessions. The shaded area (in yellow) covers the out-of-sample period, from January 2007 to March 2009.

Table C.4

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

$$C_j(t)\,\bar{s}^2(\boldsymbol{y}(t)) \equiv \frac{\hat{s}_j^2(\hat{\alpha}_j + \hat{\beta}_j y_j(t))}{\sigma^2(t)}, \ j = 1, 2, 3, \quad \text{and} \quad C_G(t) \equiv \frac{\sigma_G^2}{\sigma^2(t)},$$

where $\bar{s}(\boldsymbol{y}) \equiv \hat{s}_0 + \sum_{j=1}^{3} \hat{s}_j y_j$, $y_3(t) \equiv -\overline{\text{Sent}}_t$, and $\overline{\text{Sent}}_t$ is the UMCSENT index, rescaled as in Eq. (A.36). The sample covers monthly data for the period from January 1978 to December 2006. Panel B reports statistics for the risk-neutral counterparts to the average paths of $C_j(t)$ and $C_G(t)$. The sample covers monthly data for the period from January 1990 to December 2006.

Panel A: Contributions of factors to stock volatility

	Averages	Standard deviations
Gross inflation	0.83%	0.11%
Gross growth	74.11%	3.17%
Unobserved factor	15.40%	3.78%
Secular growth	9.66%	4.92%

Panel B: Contributions of factors to the VIX Index

	Averages	Standard deviations
Gross inflation	1.37%	0.02%
Gross growth	86.80%	2.47%
Unobserved factor	9.13%	0.92%
Secular growth	2.70%	0.15%

Finally, we provide out-of-sample results pertaining to the two methodologies of this appendix to integrate out the unobserved factor. Table C.5 reports Root Mean Squared Errors (RMSE) for the two cases where the model is fed with (i) the series of the model-implied unobserved factor, $\hat{y}_{3t}$, as calculated through Eq. (28) (second column), and (ii) the series $-\overline{\text{Sent}}_t$, where $\overline{\text{Sent}}_t$ is the UMCSENT index, rescaled as in Eq. (A.36) (third column). We rescale the UMCSENT index through Eq. (A.36), using the average and standard deviation of the model-implied unobserved factor over the five years prior to the out-of-sample experiments (January 2007 to March 2009), where $E_T(\hat{y}_3) = -2.612$ and $\sigma_T(\hat{y}_3) = 2.072$, and fixing $\sigma_T(\text{Sent}) = 10.6134$, which is the standard deviation of the UMCSENT index over the period from January 1990 to December 2006. For comparison, we also report RMSE for the benchmark cases considered in the article (Section 4.2.4), namely the case of simulations of the unobserved factor (first column), and the OLS (fourth column).

As anticipated at the beginning of Section C.3, the model generates better predictions than those stemming from OLS, in all the experiments. To summarize, (i) predictions relating to realized volatility are best performed when we feed the model with the model-implied factor; (ii) predictions relating to the VIX index are the best when we utilize the UMCSENT index. While predictions based on simulations of the unobserved factors never rank first, they rank in a stable way, compared to the predictions from alternative methodologies regarding the treatment of the unobservable factor.

Table C.5

Out-of-sample statistics under different assumptions about the treatment of the unobserved factor.

RMSE for the model and OLS

	Model (simulations)	Model (Extracted factor)	Model (UMCSENT index)	OLS
Volatility	0.0478	0.0262	0.0607	0.0700
VIX Index	0.1119	0.1121	0.1034	0.1319

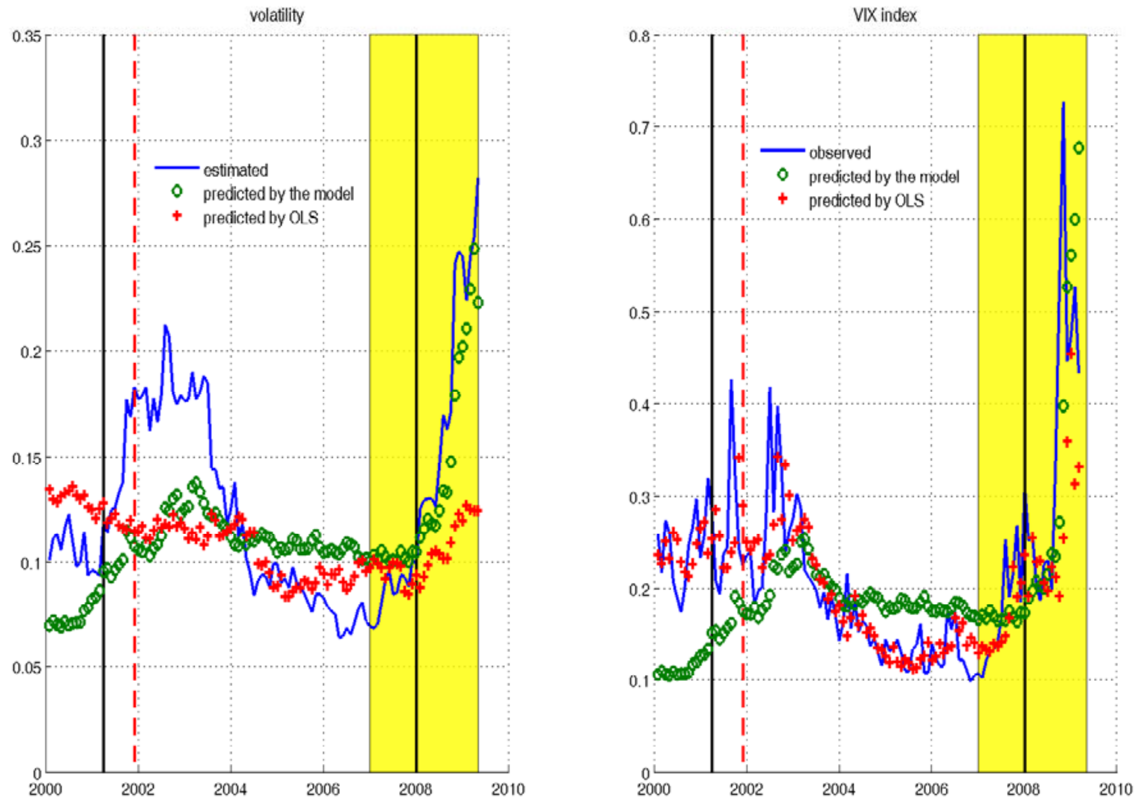

Figure C.4 – Out of sample predictions through the model-implied unobserved factor, and the subprime crisis. This figure plots one-year return volatility and the VIX index, along with its counterparts predicted by the model and by an OLS regression. The left panel depicts smoothed return volatility, defined as $\text{Vol}_t \equiv \sqrt{6\pi} \cdot 12^{-1} \sum_{i=1}^{12} |\ln (S_{t+1-i}/S_{t-i})|$, where S_t is the real stock price as of month t, along with the instantaneous standard deviation predicted by (i) the model, through Eq. (12), and (ii) the predictive part of an OLS regression of Vol_t on to past values of Vol_t, inflation and industrial production growth. The right panel depicts the VIX index, along with the VIX index predicted by (i) the model; and (ii) the predictive part of an OLS regression of the VIX index on to past values of the VIX index, inflation and industrial production growth. Each prediction is obtained by feeding the model and the predictive part of the OLS regression with the two macroeconomic factors depicted in Figure 1 (inflation and growth) and, for the model, the model-implied unobserved factor estimated as in Eq. (28), $\hat{y}_{3,t} \equiv \hat{s}_3^{-1}(\frac{S_t}{\hat{G}_t} - \hat{s}_0 - \hat{s}_1 y_{1,t} - \hat{s}_2 y_{2,t})$, where $\hat{G}_t$ is the cross-sectional average of 1000 simulations of secular growth. The sample depicted in the figure spans the period from January 2000 to March 2009. The estimation of both the model and the OLS regressions relates to the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and the vertical dashed line (in red) indicates the end of the NBER-dated recession, occurred in November 2001. The shaded area (in yellow) covers the out-of-sample period, from January 2007 to March 2009, which includes the NBER recession announced to have occurred in December 2007, and the subprime crisis, which started in June 2007.

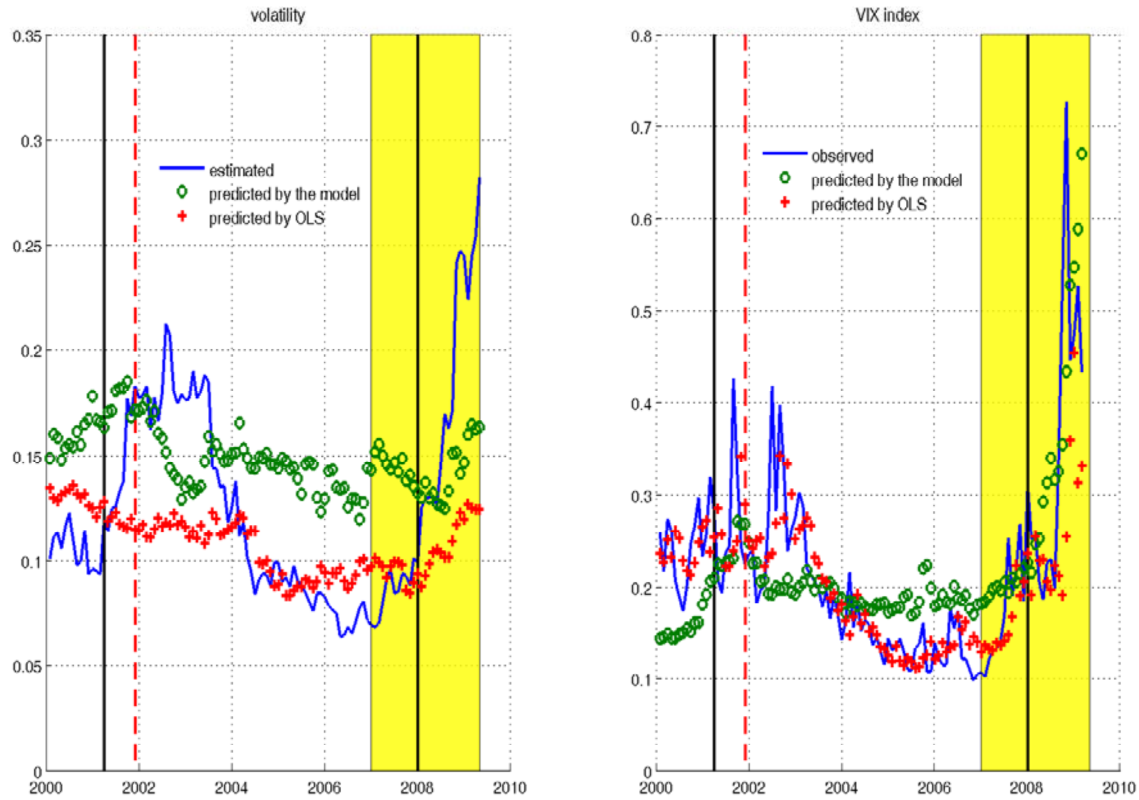

Figure C.5 – Out of sample predictions through the UMCSENT index, and the subprime crisis. This figure plots one-year return volatility and the VIX index, along with its counterparts predicted by the model and by an OLS regression. The left panel depicts smoothed return volatility, defined as $\text{Vol}_t \equiv \sqrt{6\pi} \cdot 12^{-1} \sum_{i=1}^{12} |\ln (S_{t+1-i}/S_{t-i})|$, where S_t is the real stock price as of month t, along with the instantaneous standard deviation predicted by (i) the model, through Eq. (12), and (ii) the predictive part of an OLS regression of Vol_t on to past values of Vol_t, inflation and industrial production growth. The right panel depicts the VIX index, along with the VIX index predicted by (i) the model; and (ii) the predictive part of an OLS regression of the VIX index on to past values of the VIX index, inflation and industrial production growth. Each prediction is obtained by feeding the model and the predictive part of the OLS regression with the two macroeconomic factors depicted in Figure 1 (inflation and growth) and, for the model, using $y_3(t) \equiv -\overline{\text{Sent}}_t$, where $\overline{\text{Sent}}_t$ is the UMCSENT index, rescaled as in Eq. (A.36). The sample depicted in the figure spans the period from January 2000 to March 2009. The estimation of both the model and the OLS regressions relates to the period from January 1950 to December 2006. Vertical solid lines (in black) track the beginning of NBER-dated recessions, and the vertical dashed line (in red) indicates the end of the NBER-dated recession, occurred in November 2001. The shaded area (in yellow) covers the out-of-sample period, from January 2007 to March 2009, which includes the NBER recession announced to have occurred in December 2007, and the subprime crisis, which started in June 2007.

D. References for the Supplemental material

Ang, A., Liu, J., 2004. "How to discount cashflows with time-varying expected returns. Journal of Finance 59, 2745-2783.

Angeletos, G.-M., Lorenzoni, G., Pavan, A., 2010. Beauty contests and irrational exuberance: A neoclassical approach, Working paper, MIT.

Bekaert, G., Grenadier, S.R., 2001. Stock and bond pricing in an affine economy. Working paper, Columbia Business School.

Bernanke, B.S., Gertler, M., Gilchrist, S., 1999. The financial accelerator in a quantitative business cycle framework. In: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics (Vol. 1C, Chapter 21), North-Holland Elsevier, Amsterdam, pp. 1341-1393.

Glasserman, P., Kim, K.-K., 2010. Moment explosion and stationary distributions in affine diffusion models. Mathematical Finance 20, 1-34.

Goncalves, S., White, H., 2005. Bootstrap standard error estimates for linear regression. Journal of the American Statistical Association 100, 970-979.

Hall, P., Horowitz, J. L., 1996. Bootstrap critical values for tests based on generalized-method-of-moments estimators. Econometrica 64, 891-916.

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

Mamaysky, H., 2002. A model for pricing stocks and bonds. Working paper, Yale School of Management.

Paparoditis, E., Politis, D.N., 2003. Residual-based block bootstrap for unit root testing. Econometrica 71, 813-855.

Pardoux, E., Talay, D., 1985. Discretization and simulation of stochastic differential equations. Acta Applicandae Mathematicae 3, 23-47.