Journal of Financial Economics 128 (2018) 207–233

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

An intertemporal CAPM with stochastic volatility[☆]

John Y. Campbell[a,b,*], Stefano Giglio[c,b], Christopher Polk[d], Robert Turley[e]

[a] Harvard University, Littauer Center, Cambridge MA, 02138, USA
[b] National Bureau of Economic Research, Cambridge MA, USA
[c] Yale School of Management, 165 Whitney Avenue, New Haven CT, 06511, USA
[d] London School of Economics, Department of Finance, London, WC2A 2AE, UK
[e] Dodge and Cox, 555 California St., San Francisco CA, 94104, USA

ARTICLE INFO

Article history:
Received 29 February 2016
Revised 23 January 2017
Accepted 21 February 2017
Available online 1 March 2018

JEL classification:
G11
G12

ABSTRACT

This paper studies the pricing of volatility risk using the first-order conditions of a long-term equity investor who is content to hold the aggregate equity market instead of overweighting value stocks and other equity portfolios that are attractive to short-term investors. We show that a conservative long-term investor will avoid such overweights to hedge against two types of deterioration in investment opportunities: declining expected stock returns and increasing volatility. We present novel evidence that low-frequency movements in equity volatility, tied to the default spread, are priced in the cross section of stock returns.

1. Introduction

The fundamental insight of intertemporal asset pricing theory is that long-term investors should care just as much about the returns they earn on their invested wealth as about the level of that wealth. In a simple model with a constant rate of return, for example, the sustainable level of consumption is the return on wealth multiplied by the level of wealth, and both terms in this product are equally important. In a more realistic model with time varying investment opportunities, long-term investors with relative risk aversion greater than one (conservative long-term investors) seek to hold intertemporal hedges, assets that perform well when investment opportunities deteriorate. Merton's (1973) intertemporal capital asset pricing model (ICAPM) shows that such assets should deliver lower average returns in equilibrium if they are priced from conservative long-term investors' first-order conditions.

Investment opportunities in the stock market can deteriorate either because expected stock returns decline or because the volatility of stock returns increases. The relative importance of these two types of intertemporal risk is an empirical question. In this paper, we estimate an econometric model of stock returns that captures time variation in both expected returns and volatility and permits tractable analysis of long-term portfolio choice. The model is a vector autoregression (VAR) for aggregate stock returns, realized variance, and state variables, restricted to have scalar affine stochastic volatility so that the volatilities of all shocks move proportionally.

Using this model and the first-order conditions of an infinitely lived investor with Epstein and Zin (1989, 1991) preferences, who is assumed to hold an aggregate stock index, we calculate the risk aversion needed to make the investor content to hold the market index instead of overweighting value stocks that offer higher average returns. We find that a moderate level of risk aversion, around 7,

☆ We are grateful to Torben Andersen, Gurdip Bakshi, John Cochrane, Bjørn Eraker, Bryan Kelly, Ian Martin, Sydney Ludvigson, Monika Piazzesi, Ken Singleton, Tuomo Vuolteenaho, and seminar participants at various venues for comments. We thank Josh Coval, Ken French, Nick Roussanov, Mila Getmansky Sherman, and Tyler Shumway for providing data used in the analysis.

* Corresponding author at: Harvard University, Littauer Center, Cambridge MA, 02138, USA.
E-mail address: john_campbell@harvard.edu (J.Y. Campbell).

https://doi.org/10.1016/j.jfineco.2018.02.011
0304-405X/

is sufficient to dissuade the investor from a portfolio tilt toward value stocks. Growth stocks are attractive to a moderately conservative long-term investor because they hedge against both declines in expected market returns and increases in market volatility. These considerations would not be relevant for a single-period investor.

We obtain similar results for several other equity portfolio tilts, including tilts to portfolios of stocks sorted by their past betas with market returns. High-beta stocks are attractive to a conservative long-term investor because they have hedged against increases in volatility during the past 50 years. In this way, our model helps to explain the well-known puzzle that the cross-sectional reward for market beta exposure has been low in recent decades.

We also consider managed portfolios that vary equity exposure in response to state variables. The conservative long-term investor we consider would find it attractive to hold a managed portfolio that varies equity exposure in response to time variation in expected stock returns. The reason is that we estimate only a weak correlation between expected returns and volatility, so a market timing strategy does not lead to an undesired volatility exposure.

Following Merton (1973), one could interpret the conservative long-term investor we consider in this paper as a representative investor who trades freely in all asset markets. However, this interpretation has two obstacles. First, our model does not explain why such an agent would not vary equity exposure with the level of the equity premium. Borrowing constraints can fix equity exposure at 100% when they bind, but we estimate that they will not bind at all times in our historical sample. Second, the aggregate stock index we consider here may not be an adequate proxy for all wealth, a point emphasized by many papers, including Campbell (1996), Jagannathan and Wang (1996), Lettau and Ludvigson (2001), and Lustig et al. (2013).

For both these reasons, we interpret our results in microeconomic terms, as a description of the intertemporal considerations that limit the desire of conservative long-term equity investors (including institutions such as pension funds and endowments) to follow value strategies and other equity strategies with high average returns. These considerations can contribute to the explanation of cross-sectional patterns in stock returns in a general equilibrium setting with heterogeneous investors, even if they do not provide a complete explanation in themselves.

Our empirical model provides a novel description of stochastic equity volatility that is of independent interest. Our VAR system includes not only stock returns and realized variance, but also other financial indicators including the price-smoothed earnings ratio and the default spread, the yield spread of low-rated over high-rated bonds. We find low-frequency movements in volatility tied to these variables. While this phenomenon has received little attention in the literature, we argue that it is a natural outcome of investor behavior. Because risky bonds are short the option to default over long maturities, investors in those bonds incorporate information about the long-run component of volatility when they set credit spreads. Univariate volatility forecasting methods that filter only the information in past stock returns fail to extract this low-frequency component of volatility, which is of key importance to long-horizon investors who care mostly about persistent changes in their investment opportunity set.

The organization of our paper is as follows. Section 2 reviews related literature. Section 3 presents the first-order conditions of an infinitely lived Epstein-Zin investor, allowing for a specific form of stochastic volatility, and shows how they can be used to estimate preference parameters. Section 4 presents data, econometrics, and VAR estimates of the dynamic process for stock returns and realized volatility. This section shows the empirical success of our model in forecasting long-run volatility. Section 5 introduces our basic set of test assets: portfolios of stocks sorted by value, size, and estimated risk exposures from our model. This section estimates the betas of these portfolios with news about the market's future cash flows, discount rates, and volatility and the preferences of a long-term investor that best fit the cross section of excess returns on the test assets. This section also summarizes the history of the investor's marginal utility implied by our model. Section 6 considers a larger set of equity and non-equity anomalies and asks how much the model of Section 5 contributes to explaining them. Section 7 explores alternative specifications, including the model of Bansal et al. (2014), an alternative representation of our model in terms of consumption, and additional empirical implementations of our approach. Section 8 concludes. An Online Appendix to the paper provides supporting details including a battery of robustness tests.

2. Literature review

Since Merton (1973) first formulated the ICAPM, a large empirical literature has explored the relevance of intertemporal considerations for the pricing of financial assets in general and the cross-sectional pricing of stocks in particular. One strand of this literature uses the approximate accounting identity of Campbell and Shiller (1988b) and the first-order conditions of an infinitely lived investor with Epstein-Zin preferences to obtain approximate closed-form solutions for the ICAPM's risk prices (Campbell, 1993). These solutions can be implemented empirically if they are combined with vector autoregressive estimates of asset return dynamics. Campbell and Vuolteenaho (2004), Campbell et al. (2010), and Campbell et al. (2013) use this approach to argue that value stocks outperform growth stocks on average because growth stocks hedge long-term investors against declines in the expected return on the aggregate stock market.

A weakness of these papers is that they ignore the time variation in the volatility of stock returns that is evident in the data. We remedy this weakness by augmenting the VAR system with a scalar affine stochastic volatility model in which a single state variable governs the volatility of all shocks to the VAR. Because the volatility of the volatility process itself decreases as volatility approaches zero, this specification reduces the probability that the volatility becomes negative compared with a homoskedastic volatility process, especially as the sampling frequency increases. We explore this advantage of our specification via simulations in the Online Appendix.[1] We extend the approximate closed-form ICAPM to allow for this type of stochastic volatility, and we derive three priced risk factors corresponding to three important attributes of aggregate market returns: revisions in expected future cash flows, discount rates, and volatility.

An attractive feature of our model is that the prices of these three risk factors depend on only one free parameter, the long-horizon investor's coefficient of risk aversion. This feature protects our empirical analysis from the critique of Daniel and Titman (1997, 2012) and Lewellen et al. (2010) that models with multiple free parameters can spuriously fit the returns to a set of test assets with a low-order factor structure. Our use of risk-sorted test assets further protects us from this critique.

Our work is complementary to recent research on the long-run risk model of asset prices (Bansal and Yaron, 2004), which can be traced back to insights in Kandel and Stambaugh (1991). Both the approximate closed-form ICAPM and the long-run risk model start with the first-order conditions of an infinitely lived Epstein-Zin investor. As originally stated by Epstein and Zin (1989), these first-order conditions involve both aggregate consumption growth and the return on the market portfolio of aggregate wealth. Campbell (1993) points out that the intertemporal budget constraint could be used to substitute out consumption growth, turning the model into a Merton-style ICAPM. Restoy and Weil (1998, 2011) use the same logic to substitute out the market portfolio return, turning the model into a generalized consumption capital asset pricing model (CAPM) in the style of Breeden (1979). Bansal and Yaron (2004) add stochastic volatility to the Restoy-Weil model, and subsequent theoretical and empirical research in the long-run risk framework increasingly emphasizes the importance of stochastic volatility (Bansal et al., 2012; Beeler and Campbell, 2012; Hansen, 2012). In this paper, we give the approximate closed-form ICAPM the same ability to handle stochastic volatility that its cousin, the long-run risk model, already possesses.[2]

BKSY (2014), a paper written contemporaneously with the first version of this paper, explores the effects of stochastic volatility in the long-run risk model. Like us, they find stochastic volatility to be an important feature of the time series of equity returns. BKSY propose a different benchmark asset pricing model in which a homoskedastic process drives volatility. This homoskedastic volatility process has two disadvantages. First, volatility becomes negative more frequently than when volatility follows a heteroskedastic process of the sort we assume. Second, BKSY's asset pricing solution under homoskedasticity requires an additional assumption about the covariance of news terms that is not supported by the data. The different modeling assumptions and several differences in empirical implementation account for our contrasting empirical results: BKSY estimate that volatility risk has little impact on cross-sectional risk premia and that a value-minus-growth bet has a positive beta while the aggregate stock market has a negative beta with volatility news. We find that volatility risk is very important in explaining the cross section of stock returns, that a value-minus-growth portfolio always has a negative beta with volatility news, and that the aggregate stock market's volatility beta has changed sign from negative to positive in recent decades. Section 7 presents a detailed comparison of our results with those of BKSY.

Stochastic volatility has been explored in other branches of the finance literature that we summarize in the Online Appendix. Most obviously, stochastic volatility is a prime concern of the field of financial econometrics. However, the focus has mostly been on univariate models, such as the generalized autoregressive conditional heteroskedasticity (GARCH) class of models (Engle, 1982; Bollerslev, 1986), or univariate filtering methods that use realized high-frequency volatility (Barndorff-Nielsen and Shephard, 2002; Andersen et al., 2003). A much smaller literature has, like us, looked directly at the information in other economic and financial variables concerning future volatility (Schwert, 1989; Christiansen et al., 2012; Paye, 2012; Engle et al., 2013).

3. An intertemporal model with stochastic volatility

In this section, we derive an expression for the log stochastic discount factor (SDF) of the intertemporal CAPM that allows for stochastic volatility. We then discuss the properties of the model, including the requirements for a solution to exist, the implications for asset pricing, and methods for estimation.

3.1. The stochastic discount factor

We begin by deriving the log SDF of the ICAPM with stochastic volatility.

3.1.1. Preferences

We consider an investor with Epstein–Zin preferences and write the investor's value function as

$$V_t = \left[(1-\delta) C_t^{\frac{1-\gamma}{\theta}} + \delta \left(\mathrm{E}_t \left[V_{t+1}^{1-\gamma} \right] \right)^{1/\theta} \right]^{\frac{\theta}{1-\gamma}}, \tag{1}$$

where C_t is consumption and the preference parameters are the discount factor δ, risk aversion γ, and the elasticity of intertemporal substitution (EIS) ψ. For convenience, we define $\theta = (1-\gamma)/(1-1/\psi)$.

The corresponding stochastic discount factor can be written as

$$M_{t+1} = \left(\delta \left(\frac{C_t}{C_{t+1}} \right)^{1/\psi} \right)^{\theta} \left(\frac{W_t - C_t}{W_{t+1}} \right)^{1-\theta}, \tag{2}$$

[1] Affine stochastic volatility models date back at least to Heston (1993) in continuous time. Similar models have been applied in the long-run risk literature by Eraker (2008) and Hansen (2012), among others. A continuous-time affine stochastic volatility process is guaranteed to remain positive if the drift is always positive at zero volatility, which is the case in a univariate specification. Our stochastic volatility process can go negative, albeit with low probability, because our richer multivariate specification allows the drift to be negative at zero volatility for certain configurations of the state variables.

[2] Two unpublished papers by Chen (2003) and Sohn (2010) also attempt to do this. As we discuss in detail in the Online Appendix, these papers make strong assumptions about the covariance structure of various news terms when deriving their pricing equations.

where W_t is the market value of the consumption stream owned by the agent, including current consumption C_t.[3]

We study risk premia and are therefore concerned with innovations in the SDF. We also assume that asset returns and the SDF are conditionally jointly lognormally distributed. Because we allow for changing conditional moments, we are careful to write both first and second moments with time subscripts to indicate that they can vary over time. Defining the log return on wealth $r_{t+1} = \ln(W_{t+1}/(W_t - C_t))$ and the log consumption-wealth ratio $h_{t+1} = \ln(W_{t+1}/C_{t+1})$ (denoted by h because this is the variable that determines intertemporal hedging demand), we can write the innovation in the log SDF as

$$\begin{aligned} m_{t+1} &- \mathrm{E}_t m_{t+1} \\ &= -\frac{\theta}{\psi}(\Delta c_{t+1} - \mathrm{E}_t\Delta c_{t+1}) + (\theta - 1)(r_{t+1} - \mathrm{E}_t r_{t+1}) \\ &= \frac{\theta}{\psi}(h_{t+1} - \mathrm{E}_t h_{t+1}) - \gamma(r_{t+1} - \mathrm{E}_t r_{t+1}). \end{aligned} \quad (3)$$

The second equality uses the identity $r_{t+1} - \mathrm{E}_t r_{t+1} = (\Delta c_{t+1} - \mathrm{E}_t\Delta c_{t+1}) + (h_{t+1} - \mathrm{E}_t h_{t+1})$ to substitute consumption out of the SDF, replacing it with the wealth-consumption ratio and the log return on the wealth portfolio.

3.1.2. Solving the SDF forward

The Online Appendix shows that by using Eq. (3) to price the wealth portfolio and by taking a loglinear approximation of the wealth portfolio return (that is perfectly accurate when the elasticity of intertemporal substitution equals one), we obtain a difference equation for the innovation in h_{t+1} that can be solved forward to an infinite horizon:

$$\begin{aligned} h_{t+1} &- \mathrm{E}_t h_{t+1} \\ &= (\psi - 1)(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j r_{t+1+j} \\ &\quad + \frac{1}{2}\frac{\psi}{\theta}(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j \mathrm{Var}_{t+j}\left[m_{t+1+j} + r_{t+1+j}\right] \\ &= (\psi - 1)N_{DR,t+1} + \frac{1}{2}\frac{\psi}{\theta}N_{RISK,t+1}, \end{aligned} \quad (4)$$

where ρ is a parameter of loglinearization related to the average consumption-wealth ratio that is somewhat less than one. The second equality in Eq. (4) follows CV (2004) and uses the notation N_{DR} (news about discount rates) for revisions in expected future returns. In a similar spirit, we write revisions in expectations of future risk (the variance of the future log return plus the log stochastic discount factor) as N_{RISK}.

Substituting Eq. (4) into Eq. (3) and simplifying, we obtain:

$$\begin{aligned} m_{t+1} &- \mathrm{E}_t m_{t+1} \\ &= -\gamma[r_{t+1} - \mathrm{E}_t r_{t+1}] - (\gamma - 1)N_{DR,t+1} + \frac{1}{2}N_{RISK,t+1} \\ &= -\gamma N_{CF,t+1} - [-N_{DR,t+1}] + \frac{1}{2}N_{RISK,t+1}. \end{aligned} \quad (5)$$

The first equality in Eq. (5) expresses the log SDF in terms of the market return and news about future variables. It identifies three priced factors: the market return (with price of risk γ), negative discount-rate news (with price of risk $\gamma - 1$), and news about future risk (with price of risk $-\frac{1}{2}$). This is a heteroskedastic extension of the homoskedastic ICAPM derived by Campbell (1993), with no reference to consumption or the elasticity of intertemporal substitution ψ.[4]

The second equality rewrites the model, following CV (2004), by breaking the market return into cash-flow news and discount-rate news. Cash-flow news $N_{CF,t+1}$ is defined by $N_{CF,t+1} = r_{t+1} - \mathrm{E}_t r_{t+1} + N_{DR,t+1}$. The price of risk for cash-flow news is γ times greater than the unit price of risk for negative discount-rate news. Hence, CV call betas with cash-flow news “bad betas” and those with negative discount-rate news “good betas.” The third term in Eq. (5) shows the risk price for exposure to news about future risks and did not appear in CV's model, which assumed homoskedasticity. Not surprisingly, the model implies that an asset providing positive returns when risk expectations increase offers a lower return on average. Equivalently, the log SDF is high when future volatility is anticipated to be high.

Because the elasticity of intertemporal substitution has no effect on risk prices in our model, we do not identify this parameter and, therefore, do not face the recent critique of Epstein et al. (2014) that models with a large wedge between risk aversion and the reciprocal of the EIS imply an unrealistic willingness to pay for early resolution of uncertainty.[5] However, the EIS does influence the implied behavior of the investor's consumption, a topic we explore further in Section 7.2.

3.1.3. From news about risk to news about volatility

The risk news term $N_{RISK,t+1}$ in Eq. (5) represents news about the conditional variance of returns plus the stochastic discount factor, $\mathrm{Var}_t[m_{t+1} + r_{t+1}]$. Therefore, risk news depends on the SDF and its innovations. To close the model and derive its empirical implications, we must make assumptions concerning the nature of the data generating process for stock returns and the variance terms that allow us to solve for $\mathrm{Var}_t[m_{t+1} + r_{t+1}]$ and $N_{RISK,t+1}$.

We assume that the economy is described by a first-order VAR

$$\mathbf{x}_{t+1} = \bar{\mathbf{x}} + \mathbf{\Gamma}(\mathbf{x}_t - \bar{\mathbf{x}}) + \sigma_t \mathbf{u}_{t+1}, \quad (6)$$

[3] This notational convention is not consistent in the literature. Some authors exclude current consumption from the definition of current wealth.

[4] Campbell (1993) briefly considers the heteroskedastic case, noting that, when $\gamma = 1$, $\mathrm{Var}_t[m_{t+1} + r_{t+1}]$ is a constant. This implies that N_{RISK} does not vary over time, so the stochastic volatility term disappears. Campbell claims that the stochastic volatility term also disappears when $\psi = 1$, but this is incorrect. When limits are taken correctly, N_{RISK} does not depend on ψ (except indirectly through the loglinearization parameter, ρ).

[5] We use the standard terminology to describe the two parameters of the Epstein–Zin utility function, γ as risk aversion and ψ as the elasticity of intertemporal substitution. Garcia et al. (2006) and Hansen et al. (2007), however, point out that this interpretation perhaps is not correct when γ differs from the reciprocal of ψ.

where $\mathbf{x}_{t+1}$ is an $n\times 1$ vector of state variables that has r_{t+1} as its first element, σ^2_{t+1} as its second element, and $n-2$ other variables that help to predict the first and second moments of aggregate returns. $\bar{\mathbf{x}}$ and $\mathbf{\Gamma}$ are an $n\times 1$ vector and an $n\times n$ matrix of constant parameters, and $\mathbf{u}_{t+1}$ is a vector of shocks to the state variables normalized so that its first element has unit variance. We assume that $\mathbf{u}_{t+1}$ has a constant variance-covariance matrix $\mathbf{\Sigma}$, with element $\Sigma_{11}=1$. We also define $n\times 1$ vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, all of whose elements are zero except for a unit first element in $\mathbf{e}_1$ and second element in $\mathbf{e}_2$.

The key assumption here is that a scalar random variable, σ^2_t, equal to the conditional variance of market returns, also governs time variation in the variance of all shocks to this system. Both market returns and state variables, including variance itself, have innovations whose variances move in proportion to one another. This assumption makes the stochastic volatility process affine, as in Heston (1993), and implies that the conditional variance of returns plus the stochastic discount factor is proportional to the conditional variance of returns themselves.

Given this structure, news about discount rates can be written as

$$N_{DR,t+1}=\mathbf{e}_1'\rho\mathbf{\Gamma}(\mathbf{I}-\rho\mathbf{\Gamma})^{-1}\sigma_t\mathbf{u}_{t+1}, \tag{7}$$

while implied cash-flow news is

$$N_{CF,t+1}=\left(\mathbf{e}_1'+\mathbf{e}_1'\rho\mathbf{\Gamma}(\mathbf{I}-\rho\mathbf{\Gamma})^{-1}\right)\sigma_t\mathbf{u}_{t+1}. \tag{8}$$

Our loglinear model makes the log SDF a linear function of the state variables, so all shocks to the log SDF are proportional to σ_t, and $\mathrm{Var}_t[m_{t+1}+r_{t+1}]=\omega\sigma_t^2$ for some constant parameter ω. Our specification implies that news about risk, N_{RISK}, is proportional to news about market return variance, N_V:

$$N_{RISK,t+1}=\omega\rho\mathbf{e}_2'(\mathbf{I}-\rho\mathbf{\Gamma})^{-1}\sigma_t\mathbf{u}_{t+1}=\omega N_{V,t+1}. \tag{9}$$

The parameter ω is a nonlinear function of the coefficient of relative risk aversion γ, as well as the VAR parameters and the loglinearization coefficient ρ, but it does not depend on the elasticity of intertemporal substitution ψ except indirectly through the influence of ψ on ρ. In the Online Appendix, we show that ω solves

$$\begin{aligned}\omega\sigma_t^2 &= (1-\gamma)^2\mathrm{Var}_t[N_{CF,t+1}]\\ &\quad+\omega(1-\gamma)\mathrm{Cov}_t[N_{CF,t+1},N_{V,t+1}]+\omega^2\frac{1}{4}\mathrm{Var}_t[N_{V,t+1}].\end{aligned} \tag{10}$$

γ affects ω through two main channels. First, a higher risk aversion, given the underlying volatilities of all shocks, implies a more volatile stochastic discount factor m and, therefore, higher risk. This effect is proportional to $(1-\gamma)^2$, so it increases rapidly with γ. Second, a feedback effect exists on current risk through future risk: ω appears on the right-hand side of the equation as well. Given that in our estimation we find $\mathrm{Cov}_t\left[N_{CF,t+1},N_{V,t+1}\right]<0$, this second effect makes ω increase even faster with γ.

The quadratic Eq. (10) has two solutions, but the Online Appendix shows that one of them can be disregarded. The false solution is easily identified by its implication that ω becomes infinite as volatility shocks become small. The Online Appendix also shows how to write Eq. (10) directly in terms of the VAR parameters.

Finally, substituting Eq. (9) into Eq. (5), we obtain an empirically testable expression for the SDF innovations in the ICAPM with stochastic volatility:

$$m_{t+1}-\mathrm{E}_t m_{t+1}=-\gamma N_{CF,t+1}-[-N_{DR,t+1}]+\frac{1}{2}\omega N_{V,t+1}, \tag{11}$$

where ω solves Eq. (10).

3.2. Properties and estimation of the model

We now discuss the main properties of our model and describe our estimation method.

3.2.1. Existence of a solution

With constant volatility, our model can be solved for any level of risk aversion. However, in the presence of stochastic volatility, the model admits a solution only for values of risk aversion consistent with the existence of a real solution to the quadratic equation, Eq. (10). Given our VAR estimates of the variance and covariance terms, the Online Appendix plots ω as a function of γ and shows that a real solution for ω exists when γ lies between zero and 7.2.

The Online Appendix also shows that existence of a real solution for ω requires γ to satisfy the upper bound:

$$\gamma\le 1-\frac{1}{(\rho_n-1)\sigma_{cf}\sigma_v}, \tag{12}$$

where σ_{cf} is the standard deviation of the scaled cash-flow news $N_{CF,t+1}/\sigma_t$, σ_v is the standard deviation of the scaled variance news $N_{V,t+1}/\sigma_t$, and ρ_n is the correlation between these two scaled news terms.

To develop the intuition behind these equations further, the Online Appendix studies a simple example in which the link between the existence to a solution for Eq. (10) and the existence of a value function for the representative agent can be shown analytically. The example assumes $\psi=1$, as we can then solve directly for the value function without any need for a loglinear approximation of the return on the wealth portfolio (Tallarini, 2000; Hansen et al., 2008). We find in the example that the condition for the existence of the value function coincides precisely with the condition for the existence of a real solution to the quadratic equation for ω. This result shows that the possible nonexistence of a solution to the quadratic equation for ω is a deep feature of the model, not an artifact of our loglinear approximation to the wealth portfolio return, which is not needed in the special case in which $\psi=1$. The problem arises because the value function becomes ever more sensitive to volatility as the volatility of the value function increases, and this sensitivity feeds back into the volatility of the value function, further increasing it. When this positive feedback becomes too powerful, the value function ceases to exist.[6]

[6] In the Online Appendix, we show that existence of the solution for ω also imposes a lower bound on γ: $\gamma\ge 1-(1/(\rho_n+1)\sigma_{cf}\sigma_v)$. We do not focus on this lower bound on γ because in our case, it lies far below zero, at −6.8.

In our empirical analysis, we take seriously the constraint implied by the quadratic equation, Eq. (10), and require that our parameter estimates satisfy this constraint. As a consequence, given the high average returns to risky assets in historical data, our estimate of risk aversion is often close to the estimated upper bound of 7.2.

3.2.2. Asset pricing equation and risk premia

To explore the implications of the model for risk premia, we use the general asset pricing equation under conditional lognormality,

$$\begin{aligned} 0 &= \ln \mathrm{E}_t \exp\{m_{t+1} + r_{i,t+1}\} \\ &= \mathrm{E}_t[m_{t+1} + r_{i,t+1}] + \frac{1}{2}\mathrm{Var}_t[m_{t+1} + r_{i,t+1}]. \end{aligned} \tag{13}$$

Combining this with the approximation

$$\mathrm{E}_t r_{i,t+1} + \frac{1}{2}\sigma_{it}^2 \simeq (\mathrm{E}_t R_{i,t+1} - 1), \tag{14}$$

which links expected log returns (adjusted by one-half their variance) to expected gross simple returns $R_{i,t+1}$, and subtracting Eq. (13) for any reference asset j (which could be but does not need to be a true risk-free rate) from the equation for asset i, we can write a moment condition describing the relative risk premium of i relative to j as

$$\begin{aligned} &\mathrm{E}_t\left[R_{i,t+1} - R_{j,t+1} + (r_{i,t+1} - r_{j,t+1})(m_{t+1} - \mathrm{E}_t m_{t+1})\right] \\ &= \mathrm{E}_t\Big[R_{i,t+1} - R_{j,t+1} - (r_{i,t+1} - r_{j,t+1})(\gamma N_{CF,t+1} \\ &\quad + [-N_{DR,t+1}] - \frac{1}{2}\omega N_{V,t+1})\Big] = 0, \end{aligned} \tag{15}$$

where the second equality uses Eq. (11). This expression is our main pricing equation, containing all conditional implications of the model for any pair of assets i and j. In general, the model does not restrict the covariances between the various assets' returns and the news terms. These are measured in the data and not derived from the theory (with the exception of the market portfolio itself).

We can alternatively write the moment conditions in covariance form:

$$\begin{aligned} &\mathrm{E}_t\left[R_{i,t+1} - R_{j,t+1}\right] \\ &= \gamma \mathrm{Cov}_t\left[r_{i,t+1} - r_{j,t+1}, N_{CF,t+1}\right] \\ &\quad + \mathrm{Cov}_t\left[r_{i,t+1} - r_{j,t+1}, -N_{DR,t+1}\right] \\ &\quad - \frac{1}{2}\omega \mathrm{Cov}_t\left[r_{i,t+1} - r_{j,t+1}, N_{V,t+1}\right]. \end{aligned} \tag{16}$$

As in CV (2004), this equation breaks an asset's overall covariance with unexpected returns on the wealth portfolio, $r_{t+1} - E_t r_{t+1} = N_{CF,t+1} - N_{DR,t+1}$, into two pieces, the first of which has a higher risk price than the second whenever $\gamma > 1$. Importantly, it also adds a third term capturing the asset's covariance with shocks to long-run expected future volatility.

3.2.3. Conditional and unconditional implications of the model

The moment condition in Eq. (15) summarizes the conditional asset pricing implications of the model. That expression can be conditioned down to obtain the model's unconditional implications, replacing the conditional expectation in Eq. (15) with an unconditional expectation.

A special conditional implication of the model can be obtained when we focus on the wealth portfolio and the real risk-free interest rate R_f. In this case, because both r_{t+1} and m_{t+1} are linear functions of the VAR state vector, their conditional covariance is proportional to the stochastic variance term σ_t^2:

$$\mathrm{E}_t\left[R_{t+1} - R_{f,t+1}\right] = -\mathrm{Cov}_t[r_{t+1}, m_{t+1}] \propto \sigma_t^2. \tag{17}$$

The model implies that the risk premium on the market over a risk-free real asset varies in proportion with the one-period conditional variance of the market.

This conditional restriction has some implications for the relation between news terms, in particular N_{DR} and N_V. While the restriction does not tie the two terms precisely together (because N_{DR} also reflects news about the risk-free rate), it suggests that the two should be highly correlated unless the risk-free rate is highly variable. In the special case in which the risk-free rate is constant, the model predicts $N_{DR,t+1} \propto N_{V,t+1}$.

For several reasons, we, like BKSY (2014), do not impose the conditional restriction Eq. (17) on the VAR. We want, methodologically, to let the data speak about the dynamics of returns and risks. Although imposing Eq. (17) could improve efficiency if the market is priced exactly in line with our model, our estimates would be distorted if our model is misspecified.[7]

We do not assume that we observe the riskless real return R_{t+1}^f. The standard empirical proxy, the nominal Treasury bill return, is not riskless in real terms, and recent papers have argued that this return is affected by the special liquidity of a Treasury bill, which makes it near-money (Krishnamurthy and Vissing-Jørgensen, 2012; Nagel, 2016). Such a pricing distortion implies that no model of risk and return correctly prices Treasury bills in relation to equities. Consistent with this, a large empirical literature has already rejected the restriction Eq. (17) on equity and Treasury bill returns (Campbell, 1987; Harvey, 1989, 1991; Lettau and Ludvigson, 2010), and we find that our empirical measure of σ_t^2, *EVAR*, does not significantly forecast aggregate stock returns in our unrestricted VAR.

Even though we do not impose the conditional restriction Eq. (17) on the VAR, in our empirical analysis we do test conditional asset pricing implications of the model by performing our generalized method of moments (GMM) estimation using as instruments conditioning variables implied by the model (specifically, σ_t^2). We also include a Treasury bill in the set of test assets so that we can evaluate the severity of Treasury bill mispricing relative to our model.

3.2.4. Estimation

Estimation via GMM is straightforward in this model given the moment representation of asset pricing Eq. (15).

[7] A related but distinct modeling choice is that, by contrast with BKSY (2014), we do not use ICAPM restrictions on unconditional test asset returns in estimating our VAR system. Such restrictions involve a similar trade-off between efficiency if the model is correctly specified and bias if it is misspecified. In earlier work on the two-beta ICAPM, we found that using moment conditions implied by unconditional ICAPM restrictions to estimate a VAR model is computationally challenging and can lead to numerical instability (Campbell et al., 2013).

Conditional on the news terms, the model is a linear factor model (with the caveat that both level and log returns appear), which is easy to estimate via GMM even though it imposes nonlinear restrictions on the factor risk prices. The model has only one free parameter, γ, that determines the risk prices as γ for N_{CF}, 1 for $-N_{DR}$, and $-\omega(\gamma)/2$ for N_V, where $\omega(\gamma)$ is the solution of quadratic Eq. (10) corresponding to γ and the estimated news terms.

We estimate the VAR parameters and the news terms separately via ordinary least squares (OLS) and we use GMM to estimate the preference parameter γ. Thus, our GMM standard errors for γ condition on the estimated news terms. In theory, estimating both the dynamics and the moment conditions via GMM is possible in one step. However, as discussed in CGP (2013), this estimation is involved and numerically unstable given the large number of parameters.

The moment condition Eq. (15) holds for any two assets i and j. If an inflation-indexed Treasury bill were available (whose return we would refer to as R_f), it would be a conventional choice for the reference asset j. In our empirical analysis, we use the value-weighted market portfolio as the reference asset. This is a natural choice for the reference asset as it is the portfolio that our long-term investor is assumed to hold. We also include a nominal Treasury bill return as a test asset.

Finally, we perform our GMM estimation using a prespecified diagonal weighting matrix W whose elements are the inverse of the variances of the excess returns of the test assets over the market portfolio. This approach ensures that the GMM estimation is not focusing on some extreme linear combination of the assets, while still taking into account the different variances of individual moment conditions. We have repeated our analysis using one-step and two-step efficient estimation, and the qualitative results in the paper continue to hold in these cases.

4. Predicting aggregate stock returns and volatility

In this section, we present our estimates of the VAR and discuss its predictive ability for returns and volatility.

4.1. State variables

Our full VAR specification of the vector $\mathbf{x}_{t+1}$ contains six state variables, four of which are among the five variables in CGP (2013). To those four variables, we add the Treasury bill rate R_{Tbill} (using it instead of the term yield spread used by CGP) and an estimate of conditional volatility.[8] The data are all quarterly, from 1926:2 to 2011:4.

The first variable in the VAR is the log real return on the market, r_M, the difference between the log return on the Center for Research in Security Prices (CRSP) value-weighted stock index and the log return on the Consumer Price Index. This portfolio is a standard proxy for the aggregate wealth portfolio. In the Online Appendix, we consider alternative proxies that delever the market return by combining it in various proportions with Treasury bills.

The second variable is expected market variance ($EVAR$), which is meant to capture the variance of market returns, σ_t^2, conditional on information available at time t, so that innovations to this variable can be mapped to N_V. To construct $EVAR_t$, we first create a series of within-quarter realized variance of daily returns for each time t, $RVAR_t$. We then run a regression of $RVAR_{t+1}$ on lagged realized variance ($RVAR_t$) as well as the other five state variables at time t. This regression generates a series of predicted values for $RVAR$ at each time $t+1$, which depend on information available at time t: $\widehat{RVAR}_{t+1}$. Finally, we define our expected variance at time t to be exactly this predicted value at $t+1$:

$$EVAR_t \equiv \widehat{RVAR}_{t+1}. \tag{18}$$

Although we describe our methodology in a two-step fashion in which we first estimate $EVAR$ and then use $EVAR$ in a VAR, this is only for interpretability. This approach to modeling $EVAR$ can be considered a simple renormalization of equivalent results we would find from a VAR that included $RVAR$ directly.[9]

The third variable is the log of the Standard & Poor's (S&P) 500 price-smoothed earnings ratio (PE) adapted from Campbell and Shiller (1988a), where earnings are smoothed over ten years, as in CGP (2013). The fourth is the yield on a three-month Treasury bill (R_{Tbill}) from CRSP. The fifth is the small-stock value spread (VS), constructed as described in CGP.

The sixth and final variable is the default spread (DEF), defined as the difference between the log yield on Moody's BAA and AAA bonds, obtained from the Federal Reserve Bank of St. Louis, Missouri. We include the default spread in part because that variable is known to track time series variation in expected real returns on the market portfolio (Fama and French, 1989) and because shocks to the default spread should to some degree reflect news about aggregate default probabilities, which in turn should reflect news about the market's future cash flows and volatility.

4.2. Short-run volatility estimation

For the regression model that generates $EVAR_t$ to be consistent with a reasonable data generating process for market variance, we deviate from standard OLS in two ways. First, we constrain the regression coefficients to produce fitted values (i.e., expected market return variance) that are positive. Second, given that we explicitly consider heteroskedasticity of the innovations to our variables, we estimate this regression using weighted least squares (WLS), in which the weight of each observation pair ($RVAR_{t+1}$, $\mathbf{x}_t$) is initially based on the previous period's realized variance, $RVAR_t^{-1}$. However, to ensure that

[8] The switch from the term yield spread to the Treasury bill rate was suggested by a referee of an earlier version of this paper. With either variable, our results are qualitatively and quantitatively similar.

[9] Because we weight observations based on $RVAR$ in the first stage and then re-weight observations using $EVAR$ in the second stage, our two-stage approach in practice is not exactly the same as a one-stage approach. In the Online Appendix, we explore many different ways to estimate our VAR, including using a $RVAR$-weighted, single-step estimation approach.

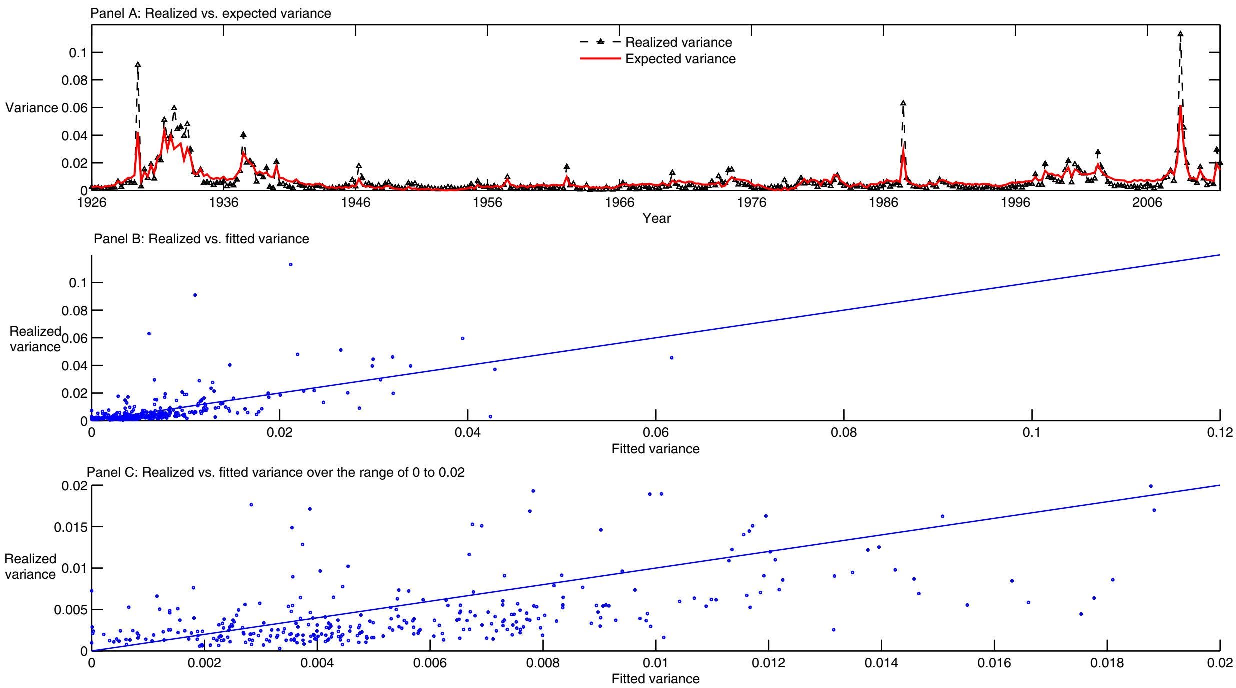

Fig. 1. Forecasting realized variance: This figure shows the results from forecasting realized variance. Panel A plots quarterly observations of realized within-quarter daily return variance over the sample period 1926:2–2011:4 and the expected variance implied by the model estimated in Table 1, Panel A. Panel B shows the full scatter plot corresponding to the regression in Table 1, Panel A. The R^2 from this regression is 38%. Panel C is similar to Panel B but zooms in on forecasts from 0 to 0.02.

the ratio of weights across observations is not extreme, we shrink these initial weights toward equal weights. We set our shrinkage factor large enough so that the ratio of the largest observation weight to the smallest observation weight is always less than or equal to five. Though admittedly somewhat ad hoc, this bound is consistent with reasonable priors on the degree of variation over time in the expected variance of market returns. More important, our results are robust to variation in this bound (see the Online Appendix). Both the constraint on the regression's fitted values and the constraint on WLS observation weights bind in the sample we study.

The first-stage regression generating the state variable $EVAR_t$ is reported in Table 1, Panel A. Not surprisingly, past realized variance strongly predicts future realized variance. More important, the regression shows that an increase in either *PE* or *DEF* predicts higher future realized volatility. Both of these results are strongly statistically significant and are a novel finding of the paper. The predictive power of very persistent variables such as *PE* and *DEF* indicates a potentially important role for lower-frequency movements in stochastic volatility.

We argue that these empirical patterns are sensible. Investors in risky bonds incorporate their expectation of future volatility when they set credit spreads, as risky bonds are short the option to default. Therefore, we expect higher *DEF* to predict higher *RVAR*. The positive predictive relation between *PE* and *RVAR* can seem surprising at first, but one has to remember that the coefficient indicates the effect of a change in *PE* holding constant the other variables, in particular the default spread *DEF*. Because the default spread should also generally depend on the equity premium and because most of the variation in *PE* is due to variation in the equity premium, we can regard *PE* as purging *DEF* of its equity premium component to reveal more clearly its forecast of future volatility. We discuss this interpretation further in Section 4.4.

The R^2 of the variance forecasting regression is nearly 38%. We illustrate this fit in several ways in Fig. 1. Panel A of the figure shows the movements of $RVAR_t$ and $EVAR_t$ over time (both variables plotted at time t), illustrating their common low-frequency variation. This graphic also highlights occasional spikes in realized variance *RVAR*, which generate high subsequent forecasts but are not themselves predicted by *EVAR*. Panel B plots the realized values at each time t, $RVAR_t$, against the forecast obtained using time $t-1$ information, $EVAR_{t-1}$, over the whole range of the data. Panel C shows the observations for which both $RVAR_t$ and $EVAR_{t-1}$ are less than 0.02 (the bottom left corner of Panel B). Fig. 1 clearly shows predictable variation in variance that is captured by our model as well as the trade-off between frequent small over-predictions of variance and infrequent large under-predictions, caused by the skewness of realized variance.

4.3. Estimation of the VAR and the news terms

In this subsection, we report our VAR estimates, then show the implied cash-flow, discount-rate, and volatility news series.

Table 1
Vector autoregression (VAR) estimation.
The table shows the weighted least squares (WLS) parameter estimates for a first-order VAR model. The state variables in the VAR are the log real return on the Center for Research in Security Prices (CRSP) value-weight index (r_M), the realized variance ($RVAR$) of within-quarter daily simple returns on the CRSP value-weight index, the log ratio of the Standard & Poor's (S&P) 500's price to the S&P 500's ten-year moving average of earnings (PE), the log three-month Treasury bill yield (r_{Tbill}), the default yield spread (DEF) in percentage points, measured as the difference between the log yield on Moody's BAA bonds and the log yield on Moody's AAA bonds, and the small-stock value spread (VS), the difference in the log book-to-market ratios of small-value and small-growth stocks. The small-value and small-growth portfolios are two of the six elementary portfolios constructed by Davis et al. (2000). For the sake of interpretation, we estimate the VAR in two stages. Panel A reports the WLS parameter estimates of a first-stage regression forecasting $RVAR$ with the VAR state variables. The forecasted values from this regression are used in the second stage of the estimation procedure as the state variable $EVAR$, replacing $RVAR$ in the second-stage VAR. Panel B reports WLS parameter estimates of the full second-stage VAR. Initial WLS weights on each observation are inversely proportional to $RVAR_t$ and $EVAR_t$ in the first and second stage, respectively, and are then shrunk to equal weights so that the maximum ratio of actual weights used is less than or equal to five. In addition, the forecasted values for both $RVAR$ and $EVAR$ are constrained to be positive. In Panels A and B, Columns 1–7 report coefficients on an intercept and the six explanatory variables, and Column 8 shows the implied R^2 statistic for the unscaled model. Bootstrapped standard errors that take into account the uncertainty in generating $EVAR$ are in parentheses. Panel C reports the correlation ("Corr/std") matrices of both the unscaled and scaled shocks from the second-stage VAR, with shock standard deviations on the diagonal. Panel D reports the results of regressions forecasting the squared second-stage residuals from the VAR with $EVAR_t$. For readability, the estimates in the regression forecasting $r_{Tbill,t+1}$ with $EVAR_t$ are multiplied by ten thousand. Bootstrap standard errors that take into account the uncertainty in generating $EVAR$ are in parentheses. The sample period for the dependent variables is 1926:3–2011:4, with 342 quarterly data points.

Panel A: Forecasting quarterly realized variance ($RVAR_{t+1}$)

Constant (1)	$r_{M,t}$ (2)	$RVAR_t$ (3)	PE_t (4)	$r_{Tbill,t}$ (5)	DEF_t (6)	VS_t (7)	R^2 (8)
−0.020	−0.005	0.374	0.006	−0.042	0.006	0.000	37.80%
(0.009)	(0.005)	(0.066)	(0.002)	(0.057)	(0.001)	(0.003)	

Panel B: VAR estimates

Second stage	Constant (1)	$r_{M,t}$ (2)	$EVAR_t$ (3)	PE_t (4)	$r_{Tbill,t}$ (5)	DEF_t (6)	VS_t (7)	R^2 (8)
$r_{M,t+1}$	0.221	0.041	0.335	−0.042	−0.810	0.010	−0.051	3.36%
	(0.113)	(0.063)	(2.143)	(0.032)	(0.736)	(0.022)	(0.035)	
$EVAR_{t+1}$	−0.016	−0.002	0.441	0.005	−0.021	0.004	0.001	60.78%
	(0.007)	(0.001)	(0.057)	(0.002)	(0.046)	(0.001)	(0.002)	
PE_{t+1}	0.155	0.130	0.674	0.961	−0.399	−0.001	−0.024	94.29%
	(0.113)	(0.062)	(2.112)	(0.032)	(0.734)	(0.022)	(0.035)	
$r_{Tbill,t+1}$	0.001	0.002	−0.084	0.001	0.948	0.001	−0.001	94.07%
	(0.004)	(0.002)	(0.075)	(0.001)	(0.024)	(0.001)	(0.001)	
DEF_{t+1}	0.194	−0.293	11.162	−0.118	4.102	0.744	0.175	88.22%
	(0.309)	(0.176)	(5.838)	(0.086)	(1.925)	(0.062)	(0.094)	
VS_{t+1}	0.147	0.069	2.913	−0.017	−0.253	−0.004	0.932	93.93%
	(0.111)	(0.065)	(2.169)	(0.031)	(0.705)	(0.022)	(0.034)	

Panel C: Correlations and standard deviations

Corr/std	r_M	$EVAR$	PE	r_{Tbill}	DEF	VS
unscaled						
r_M	0.105	−0.509	0.907	−0.041	−0.482	−0.039
$EVAR$	−0.509	0.004	−0.592	−0.163	0.688	0.106
PE	0.907	−0.592	0.099	−0.004	−0.598	−0.066
r_{Tbill}	−0.041	−0.163	−0.004	0.003	−0.111	0.013
DEF	−0.482	0.688	−0.598	−0.111	0.287	0.323
VS	-0.039	0.106	-0.066	0.013	0.323	0.086
scaled						
r_M	1.138	-0.494	0.905	-0.055	-0.367	0.022
$EVAR$	-0.494	0.044	-0.570	-0.178	0.664	0.068
PE	0.905	-0.570	1.047	-0.014	-0.479	0.005
r_{Tbill}	-0.055	-0.178	−0.014	0.041	-0.160	-0.001
DEF	-0.367	0.664	-0.479	-0.160	2.695	0.273
VS	0.022	0.068	0.005	-0.001	0.273	0.996

Panel D: Heteroskedastic shocks

Squared, second-stage, unscaled residual	Constant	$EVAR_t$	R^2
$r_{M,t+1}$	-0.002	1.85	20.43%
	(0.003)	(0.283)	
$EVAR_{t+1}$	0.000	0.004	6.36%
	(0.000)	(0.001)	
PE_{t+1}	-0.004	1.89372	19.75%
	(0.003)	(0.289)	
$r_{Tbill,t+1}$	0.111	0.283	-0.29%
	(0.054)	(4.542)	
DEF_{t+1}	-0.113	27.166	27.50%
	(0.041)	(3.411)	
VS_{t+1}	0.004	0.472	5.57%
	(0.002)	(0.133)	

4.3.1. VAR estimates

We estimate a first-order VAR as in Eq. (6), where $\mathbf{x}_{t+1}$ is a 6×1 vector of state variables ordered as follows:

$$\mathbf{x}_{t+1} = [r_{M,t+1} \; EVAR_{t+1} \; PE_{t+1} \; R_{Tbill,t+1} \; DEF_{t+1} \; VS_{t+1}], \quad (19)$$

so that the real market return $r_{M,t+1}$ is the first element and *EVAR* is the second element. $\bar{\mathbf{x}}$ is a 6×1 vector of the means of the variables, and $\boldsymbol{\Gamma}$ is a 6×6 matrix of constant parameters. Finally, $\sigma_t \mathbf{u}_{t+1}$ is a 6×1 vector of innovations, with the conditional variance-covariance matrix of $\mathbf{u}_{t+1}$ a constant $\boldsymbol{\Sigma}$, so that the parameter σ_t^2 scales the entire variance-covariance matrix of the vector of innovations.

The first-stage regression forecasting realized market return variance described in Section 4.2 generates the variable *EVAR*. The theory in Section 3 assumes that σ_t^2, proxied for by *EVAR*, scales the variance-covariance matrix of state variable shocks. Thus, as in the first stage, we estimate the second-stage VAR using WLS, in which the weight of each observation pair $(\mathbf{x}_{t+1}, \mathbf{x}_t)$ is initially based on $(EVAR_t)^{-1}$. We continue to constrain both the weights across observations and the fitted values of the regression forecasting *EVAR*.

Table 1, Panel B, presents the results of the VAR estimation for the full sample (1926:2 to 2011:4).[10] We report bootstrap standard errors for the parameter estimates of the VAR that take into account the uncertainty generated by forecasting variance in the first stage. Consistent with previous research, we find that *PE* negatively predicts future returns, though the *t*-statistic indicates only marginal significance. The value spread has a negative but not statistically significant effect on future returns. In our specification, a higher conditional variance, *EVAR*, is associated with higher future returns, though the effect is not statistically significant. The relatively high degree of correlation among *PE*, *DEF*, *VS*, and *EVAR* complicates the interpretation of the individual effects of those variables. As for the other novel aspects of the transition matrix, both high *PE* and high *DEF* predict higher future conditional variance of returns. High past market returns forecast lower *EVAR*, higher *PE*, and lower *DEF*.[11]

Panel C of Table 1 reports the sample correlation matrices of both the unscaled residuals $\sigma_t \mathbf{u}_{t+1}$ and the scaled residuals $\mathbf{u}_{t+1}$. The correlation matrices report standard deviations on the diagonals. A comparison of the standard deviations of the unscaled and scaled market return residuals provides a rough indication of the effectiveness of our empirical solution to the heteroskedasticity of the VAR. The scaled return residuals should have unit standard deviation, and our implementation results in a sample standard deviation of 1.14.[12]

Panel D reports the coefficients of a regression of the squared unscaled residuals $\sigma_t \mathbf{u}_{t+1}$ of each VAR equation on a constant and *EVAR*. These results are broadly consistent with our assumption that *EVAR* captures the conditional volatility of the market return and other state variables. The coefficient on *EVAR* in the regression forecasting the squared market return residuals is 1.85, not the theoretically expected value of one, but this coefficient is sensitive to the weighting scheme used in the regression. We can reject the null hypothesis that all six regression coefficients are jointly zero or negative. This evidence is consistent with the volatilities of all innovations being driven by a common factor, as we assume, although, empirically, other factors also can influence the volatilities of certain variables.

4.3.2. News terms

Panels A and B of Table 2 present the variance-covariance matrix and the standard deviation and correlation matrix of the news terms, estimated as described above. We find, consistent with previous research, that discount-rate news is nearly twice as volatile as cash-flow news.

The interesting new results in this table concern the variance news term N_V. News about future variance has significant volatility, with nearly a third of the variability of discount-rate news. Variance news is negatively correlated (-0.12) with cash-flow news. As one could expect from the literature on the leverage effect (Black, 1976; Christie, 1982), news about low cash flows is associated with news about higher future volatility. N_V is close to uncorrelated (-0.03) with discount-rate news.[13] The net effect of these correlations, in Panel C of Table 2, is a correlation close to zero (again -0.03) between our measure of volatility news and contemporaneous market returns.

Panel D of Table 2 reports the decomposition of the vector of innovations $\sigma_t \mathbf{u}_{t+1}$ into the three terms $N_{CF,t+1}, N_{DR,t+1}$, and $N_{V,t+1}$. As shocks to *EVAR* are just a linear combination of shocks to the underlying state variables, which includes *RVAR*, we unpack *EVAR* to express the news terms as a function of r_M, *PE*, R_{Tbill}, *VS*, *DEF*, and *RVAR*. The panel shows that innovations to *RVAR* are mapped more than one-to-one to news about future volatility. However, several of the other state variables also drive news about volatility. Specifically, innovations in *PE*, *DEF*, and *VS* are associated with news of higher future volatility. Panel D also indicates that all state variables with the exception of R_{Tbill} are statistically significant in terms of their contribution to at least one of the three news terms. We choose to leave R_{Tbill} in the VAR, though

[10] In our robustness test, we show that our findings continue to hold if we either estimate our model's news terms out-of-sample or allow the coefficients in the first two regressions of the VAR to vary across the early and modern subsamples.

[11] One worry is that many elements of the transition matrix are estimated imprecisely. Though these estimates can be zero, their nonzero but statistically insignificant in-sample point estimates, in conjunction with the highly nonlinear function that generates discount-rate and volatility news, can result in misleading estimates of risk prices. However, the Online Appendix shows that we continue to find an economically significant negative volatility beta for value-minus-growth bets if we instead employ a partial VAR in which, via a standard iterative process, only variables with *t* -statistics greater than 1.0 are included in each VAR regression.

[12] A comparison of the unscaled and scaled autocorrelation matrices, in the Online Appendix, reveals in addition that much of the sample autocorrelation in the unscaled residuals is eliminated by our WLS approach.

[13] Though the point estimate of this correlation is negative, the large standard error implies that we cannot reject the volatility feedback effect (Campbell and Hentschel, 1992; Calvet and Fisher, 2007), which generates a positive correlation. For related research, see French et al. (1987).

Table 2

Cash-flow, discount-rate, and variance news for the market portfolio.

The table shows the properties of cash-flow news (N_{CF}), discount-rate news (N_{DR}), and volatility news (N_V) implied by the vector autoregression (VAR) model of Table 1. Panel A shows the covariance matrix of the news terms. For readability, these estimates are scaled by one hundred. Panel B shows the correlation matrix of the news terms with standard deviations on the diagonal. Panel C shows the correlations of shocks to individual state variables with the news terms. Panel D shows the functions $(\mathbf{e1}' + \mathbf{e1}'\lambda_{DR},\ \mathbf{e1}'\lambda_{DR},\ \mathbf{e2}'\lambda_V)$ that map the state variable shocks to cash-flow, discount-rate, and variance news. We define $\lambda_{DR} \equiv \rho\mathbf{\Gamma}(\mathbf{I}-\rho\mathbf{\Gamma})^{-1}$ and $\lambda_V \equiv \rho(\mathbf{I}-\rho\mathbf{\Gamma})^{-1}$, where $\mathbf{\Gamma}$ is the estimated VAR transition matrix from Table 1 and ρ is set to 0.95 per annum. r_M is the log real return on the Center for Research in Security Prices (CRSP) value-weight index. *RVAR* is the realized variance of within-quarter daily simple returns on the CRSP value-weight index. *PE* is the log ratio of the Standard & Poor's (S&P) 500's price to the S&P 500's ten-year moving average of earnings. r_{Tbill} is the log three-month Treasury yield. *DEF* is the default yield spread in percentage points, measured as the difference between the log yield on Moody's BAA bonds and the log yield on Moody's AAA bonds. *VS* is the small-stock value spread, the difference in the log book-to-market ratios of small-value and small-growth stocks. Bootstrap standard errors that take into account the uncertainty in generating *EVAR* are in parentheses.

	N_{CF}	N_{DR}	N_V
Panel A: News terms			
N_{CF}	0.236	−0.018	−0.015
	(0.087)	(0.119)	(0.030)
N_{DR}	−0.018	0.838	−0.008
	(0.119)	(0.270)	(0.065)
N_V	−0.015	−0.008	0.065
	(0.030)	(0.065)	(0.030)
Panel B: Correlations and standard deviations of news terms			
N_{CF}	0.049	−0.041	−0.121
	(0.008)	(0.225)	(0.264)
N_{DR}	−0.041	0.092	−0.034
	(0.225)	(0.014)	(0.355)
N_V	−0.121	−0.034	0.025
	(0.264)	(0.355)	(0.007)
Panel C: Correlations with shocks to state variables			
r_M shock	0.497	−0.888	−0.026
	(0.213)	(0.045)	(0.332)
EVAR shock	−0.040	0.564	0.660
	(0.196)	(0.143)	(0.174)
PE shock	0.158	−0.960	−0.097
	(0.239)	(0.044)	(0.354)
r_{Tbill} shock	−0.372	−0.151	−0.034
	(0.219)	(0.142)	(0.331)
DEF shock	−0.041	0.533	0.751
	(0.188)	(0.115)	(0.223)
VS shock	−0.397	−0.165	0.567
	(0.187)	(0.141)	(0.261)
Panel D: Functions of shocks to state variables			
r_M shock	0.908	−0.092	−0.011
	(0.031)	(0.031)	(0.015)
RVAR shock	−0.300	−0.300	1.280
	(1.134)	(1.134)	(0.571)
PE shock	−0.814	−0.814	0.187
	(0.167)	(0.167)	(0.084)
r_{Tbill} shock	−4.245	−4.245	0.867
	(3.635)	(3.635)	(1.821)
DEF shock	0.008	0.008	0.079
	(0.034)	(0.034)	(0.017)
VS shock	−0.248	−0.248	0.099
	(0.127)	(0.127)	(0.064)

its presence in the system makes little difference to our conclusions.

Fig. 2 plots the N_{CF}, $-N_{DR}$, and N_V series. To emphasize lower-frequency movements and to improve the readability of the figure, we normalize each series by its standard deviation and then smooth (for plotting purposes only), using an exponentially weighted moving average with a quarterly decay parameter of 0.08. This decay parameter implies a half-life of approximately two years. The pattern of N_{CF} and $-N_{DR}$ is consistent with previous research, for example, Fig. 1 of CV (2004). As a consequence, we focus on the smoothed series for market variance news. Considerable time variation exists in N_V. We find episodes of news of high future volatility during the Great Depression and just before the beginning of World War II, followed by a period of little news until the late 1960s. From then on, periods of positive volatility news alternate with periods of negative volatility news in cycles of three to five years. Spikes in news about future volatility are found in the early 1970s (following the oil shocks), in the late 1970s and again following the 1987 stock market crash. The late 1990s are characterized by strongly negative news about future returns and at the same time, higher expected future volatility. The recession of the late 2000s is characterized by strongly negative cash-flow news, together with a spike in volatility of the highest magnitude in our sample. The recovery from the 2007–2009 financial crisis has brought positive cash-flow news together with news about lower future volatility.

4.4. Predicting long-run volatility

The predictability of volatility, and especially of its long-run component, is central to this paper. We have shown that volatility is strongly predictable, specifically by variables beyond lagged realizations of volatility itself: *PE* and *DEF* contain essential information about future volatility. We have also proposed a VAR-based methodology to construct long-horizon forecasts of volatility that incorporate all the information in lagged volatility as well as in the additional predictors such as *PE* and *DEF*.

We now ask how well our proposed long-run volatility forecast captures the long-horizon component of volatility. In the Online Appendix, we regress realized, discounted, annualized long-run variance up to period h,

$$LHRVAR_h = \frac{4\Sigma_{j=1}^{h}\rho^{j-1}RVAR_{t+j}}{\Sigma_{j=1}^{h}\rho^{j-1}}, \tag{20}$$

on the variables included in our VAR system, the VAR long-horizon forecast, and some alternative forecasts of long-run variance. We focus on a ten year horizon ($h = 40$) as longer horizons come at the cost of fewer independent observations. However, the Online Appendix confirms that our results are robust to horizons ranging from one to 15 years.

We estimate, as alternatives to the VAR approach, two standard GARCH-type models, designed to capture the long-run component of volatility: the two-component exponential (EGARCH) model proposed by Adrian and Rosenberg (2008), and the fractionally integrated (FIGARCH)

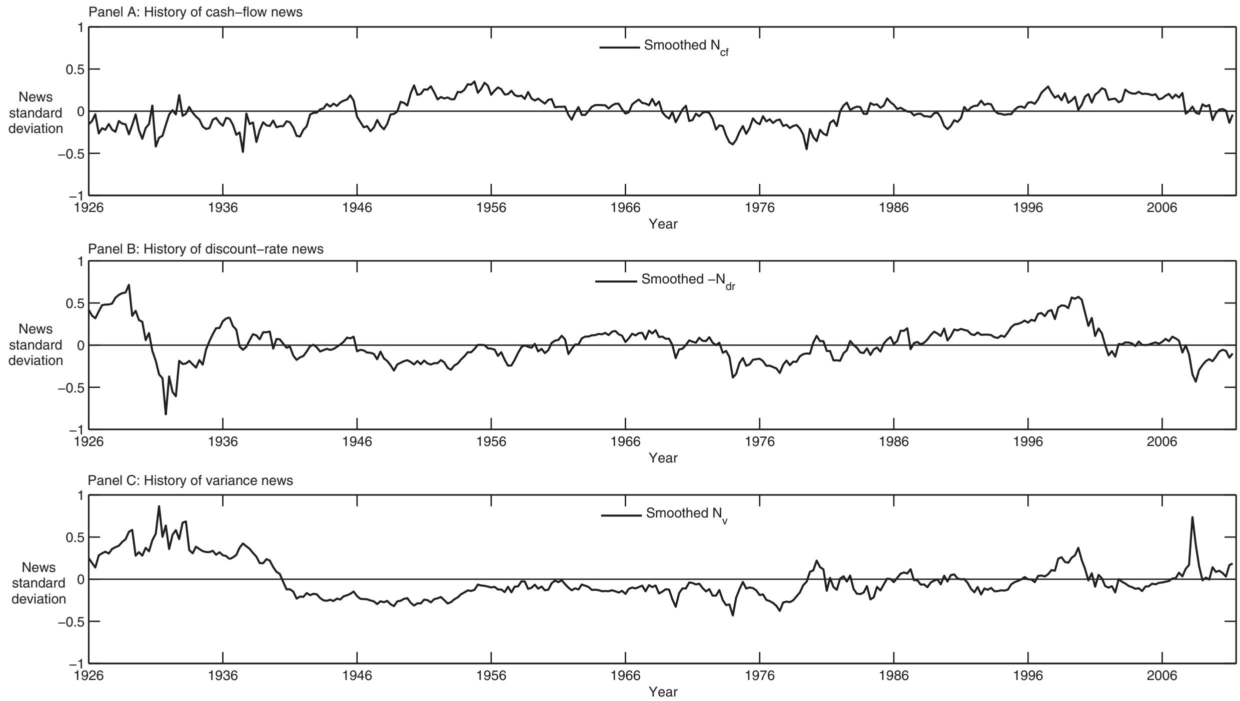

Fig. 2. The history of market news: This figure plots cash-flow news, the negative of discount-rate news, and variance news. The series are normalized by their standard deviations and then smoothed with a trailing exponentially weighted moving average in which the decay parameter is set to 0.08 per quarter, and the smoothed normalized news series is generated as $MA_t(N) = 0.08N_t + (1 - 0.08)MA_{t-1}(N)$. This decay parameter implies a half-life of two years. The sample period is 1926:2–2011:4.

model of Baillie et al. (1996). We estimate both GARCH models using the full sample of daily returns and then generate the appropriate forecast of $LHRVAR_{40}$. To these two models, we add the set of variables from our VAR and compare the forecasting ability of these different models. We find that while the EGARCH and FIGARCH forecasts do forecast long-run volatility, our VAR variables provide as good or better explanatory power, and *RVAR*, *PE* and *DEF* are strongly statistically significant. Our long-run VAR forecast has a coefficient of 1.02, which remains highly significant at 0.82 even in the presence of the FIGARCH forecast. We also find that *DEF* does not predict long-horizon volatility in the presence of our VAR forecast, implying that the VAR model captures the long-horizon information in the default spread.

The Online Appendix also examines more carefully the links between *PE*, *DEF*, and $LHRVAR_{40}$. By itself, *PE* has almost no information about low-frequency variation in volatility. In contrast, *DEF* forecasts nearly 22% of the variation in $LHRVAR_{40}$. Furthermore, if we use the component of *DEF* that is orthogonal to *PE*, which we call *DEFO* or the *PE*-adjusted default spread, the R^2 increases to over 51%. Our interpretation of these results is that *DEF* contains information about future volatility because risky bonds are short the option to default. However, *DEF* also contains information about future aggregate risk premia. We know from previous work that much of the variation in *PE* reflects aggregate risk premia. Therefore, including *PE* in the volatility forecasting regression cleans up variation in *DEF* resulting from variation in aggregate risk premia and thus sharpens the link between *DEF* and future volatility. Because *PE* and *DEF* are negatively correlated (default spreads are relatively low when the market trades rich), both *PE* and *DEF* receive positive coefficients in the multiple regression.

Fig. 3 provides a visual summary of the long-run volatility-forecasting power of our key VAR state variables and our interpretation. Panel A plots $LHRVAR_{40}$ together with lagged *DEF* and *PE*. The graph confirms the strong negative correlation between *PE* and *DEF* (correlation of -0.6) and highlights the way both variables track long-run movements in long-run volatility. To isolate the contribution of the default spread in predicting long run volatility, Panel B plots $LHRVAR_{40}$ together with *DEFO*, the *PE*-adjusted default spread that is orthogonal to the market's smoothed price-earnings ratio. The improvement in fit moving from Panel A to Panel B is clear.

The contrasting behavior of *DEF* and *DEFO* in the two panels during episodes such as the tech boom help illustrate the workings of our story. Taken in isolation, the relatively stable default spread throughout most of the late 1990s would predict little change in future market volatility. However, once the declining equity premium over that period is taken into account (as shown by the rapid increase in *PE*), a high *PE*-adjusted default spread in the late 1990s forecasted much higher volatility ahead.

As a further check on the usefulness of our VAR approach, we compare in the Online Appendix our variance forecasts to option-implied variance forecasts over the period 1998–2011. When both the VAR and option data are used to predict realized variance, the VAR forecasts drive

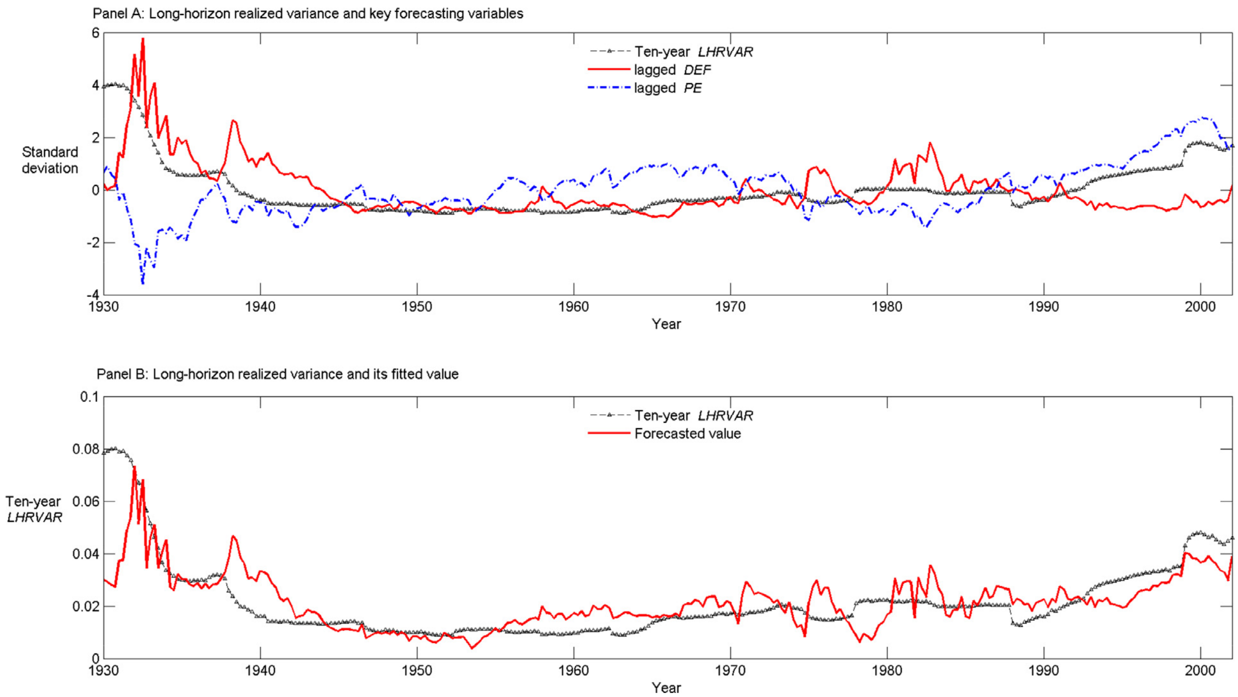

Fig. 3. Modeling low-frequency variation in realized market variance: We measure long-horizon realized variance (*LHRVAR*) as the annualized discounted sum of within-quarter daily return variance, $LHRVAR_h = \frac{4*\Sigma_{j=1}^{h}\rho^{j-1}RVAR_{t+j}}{\Sigma_{j=1}^{h}\rho^{j-1}}$. Each panel plots quarterly observations of ten-year realized variance, $LHRVAR_{40}$, over the sample period 1930:1–2001:1. In Panel A, in addition to $LHRVAR_{40}$, we plot lagged *PE* and *DEF*. In Panel B, in addition to $LHRVAR_{40}$, we plot the fitted value from a regression forecasting $LHRVAR_{40}$ with *DEFO*, defined as *DEF* orthogonalized to demeaned *PE*. The Online Appendix reports the weighted least squares estimates of this forecasting regression.

out the option-implied forecasts while remaining statistically and economically significant.

Taken together, these results make a strong case that credit spreads and valuation ratios contain information about future volatility not captured by simple univariate models, even those designed to fit long-run movements in volatility. Furthermore, our VAR method for calculating long-horizon forecasts preserves this information.

5. Estimating the ICAPM using equity portfolios sorted by size, value, and risk

In this section, we estimate and test the ICAPM with stochastic volatility using various cross sections of equity returns.

5.1. Construction of test assets

In addition to the VAR state variables, our analysis requires excess returns on a set of test assets. In this subsection, we construct several sets of equity portfolios sorted by value, size, and risk estimates from our model. Full details on the construction method are provided in the Online Appendix.

Because the long-term investor in our model is assumed to hold the equity market, we measure all excess returns relative to the market portfolio. Our primary cross section consists of the excess returns over the market on 25 portfolios sorted by size and value (ME and BE/ME), studied in Fama and French (1993), extended in Davis et al. (2000), and made available by Professor Kenneth French on his website.[14] To this cross section, we add the excess return on a Treasury bill over the market (the negative of the usual excess return on the market over a Treasury bill), which gives an initial set of 26 characteristic-sorted test assets.

We incorporate additional assets in our tests to guard against the concerns of Daniel and Titman (1997, 2012) and Lewellen et al. (2010) that characteristic-sorted portfolios can have a low-order factor structure that is easily fit by spurious models. We construct a second set of six risk-sorted portfolios, double-sorted on past multiple betas with market returns and variance innovations (approximated by a weighted average of changes in the VAR explanatory variables).

We also consider excess returns on equity portfolios that are formed based on both characteristics and past exposures to variance innovations. One possible explanation for our finding that growth stocks hedge volatility relative to value stocks is that growth firms are more likely to hold real options, whose value increases with volatility. To test this interpretation, we sort stocks based on two firm characteristics that are often used to proxy for the presence of real options and that are available for a large percentage of firms throughout our sample period: the ra-

[14] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

tio of book equity to market equity (BE/ME) and idiosyncratic volatility (*ivol*). Having formed nine portfolios using a two-way characteristic sort, we split each of these portfolios into two subsets based on pre-formation estimates of each stock's simple beta with variance innovations. One could expect that sorts on simple instead of partial betas would be more effective in establishing a link between pre-formation and post-formation estimates of volatility beta, because the market is correlated with volatility news. This gives us eighteen portfolios sorted on both characteristics and risk.

Combining all the above portfolios, we have a set of 50 test assets. We finally create managed or scaled versions of all these portfolios by interacting them with our volatility forecast *EVAR*. The managed portfolios increase their exposure to test assets at times when market variance is expected to be high. With both unscaled and scaled portfolios, we have a total of one hundred test assets.[15]

Previous research, particularly CV (2004), has found important differences in the risks of value stocks in the periods before and after 1963. Accordingly, we consider two main subsamples, which we call early (1931:3–1963:3) and modern (1963:4–2011:4). A successful model should be able to fit the cross section of test asset returns in both these periods with stable parameters.

5.2. Beta measurement

We first examine the betas implied by the covariance form of the model in Eq. (16). We cosmetically multiply and divide all three covariances by the sample variance of the unexpected log real return on the market portfolio to facilitate comparison with previous research, defining

$$\beta_{i,CF_M} \equiv \frac{Cov(r_{i,t}, N_{CF,t})}{Var(r_{M,t} - E_{t-1} r_{M,t})}, \tag{21}$$

$$\beta_{i,DR_M} \equiv \frac{Cov(r_{i,t}, -N_{DR,t})}{Var(r_{M,t} - E_{t-1} r_{M,t})}, \tag{22}$$

$$\beta_{i,V_M} \equiv \frac{Cov(r_{i,t}, N_{V,t})}{Var(r_{M,t} - E_{t-1} r_{M,t})}. \tag{23}$$

The risk prices on these betas are just the variance of the market return innovation times the risk prices in Eq. (16).

We estimate cash-flow, discount-rate, and variance betas using the fitted values of the market's cash-flow, discount-rate, and variance news estimated in Section 4. We estimate simple WLS regressions of each portfolio's log returns on each news term, weighting each time $t+1$ observation pair by the weights used to estimate the VAR in Table 1, Panel B. We then scale the regression loadings by the ratio of the sample variance of the news term in question to the sample variance of the unexpected log real return on the market portfolio to generate estimates for our three-beta model.

5.2.1. Characteristic-sorted portfolios

Table 3, Panel A, shows the estimated betas for the characteristic-sorted portfolios over the 1931–1963 period. To save space, we omit the betas for portfolios in the second and fourth quintiles of each characteristic, retaining only the first, third, and fifth quintiles. The full table can be found in the Online Appendix.

The portfolios are organized in a square matrix with growth stocks at the left, value stocks at the right, small stocks at the top, and large stocks at the bottom. At the right edge of the matrix we report the differences between the extreme growth and extreme value portfolios in each size group. Along the bottom of the matrix, we report the differences between the extreme small and extreme large portfolios in each BE/ME category. The top matrix displays post-formation cash-flow betas, the middle matrix displays post-formation discount-rate betas, and the bottom matrix displays post-formation variance betas. In square brackets after each beta estimate, we report a standard error, calculated conditional on the realizations of the news series from the aggregate VAR model.

In the pre-1963 sample period, value stocks (except those in the smallest size quintile) have both higher cash-flow and discount-rate betas than growth stocks. An equal-weighted average of the extreme value stocks across all size quintiles has a cash-flow beta 0.12 higher than an equal-weighted average of the extreme growth stocks. The average difference in estimated discount-rate betas, 0.25, is in the same direction. Similar to value stocks, small stocks have consistently higher cash-flow betas and discount-rate betas than large stocks in this sample (by 0.16 and 0.36, respectively, for an equal-weighted average of the smallest stocks across all value quintiles relative to an equal-weighted average of the largest stocks). These differences are extremely similar to those in CV (2004) despite the exclusion of the 1929–1931 subperiod, the replacement of the excess log market return with the log real return, and the use of a richer, heteroskedastic VAR.

The new finding in Table 3, Panel A, is that value stocks and small stocks are also riskier in terms of volatility betas. An equal-weighted average of the extreme value stocks across all size quintiles has a volatility beta 0.06 lower than an equal-weighted average of the extreme growth stocks. An equal-weighted average of the smallest stocks across all value quintiles has a volatility beta that is 0.06 lower than an equal-weighted average of the largest stocks. In summary, value and small stocks were unambiguously riskier than growth and large stocks over the 1931–1963 period.

Table 3, Panel B, reports the corresponding estimates for the post-1963 period. As shown in this subsample by CV (2004), value stocks still have slightly higher cash-flow betas than growth stocks, but much lower discount-rate betas. Our new finding here is that value stocks continue to have much lower volatility betas, and the spread in volatility betas is even greater than in the early period. The volatility beta for the equal-weighted average of the extreme value stocks across size quintiles is 0.11 lower than the volatility beta of an equal-weighted average of the extreme growth stocks, a difference that is more than 85% higher than the corresponding difference in the early period.

These results imply that in the post-1963 period, when the CAPM has difficulty explaining the low returns on

[15] Table OA1 in the Online Appendix reports summary statistics for these portfolios.

Table 3
Cash-flow, discount-rate, and variance betas.
The table shows the estimated cash-flow ($\widehat{\beta}_{CF}$), discount-rate ($\widehat{\beta}_{DR}$), and variance betas ($\widehat{\beta}_{V}$) for the 25 ME- and BE/ME-sorted portfolios (Panels A and B) and six risk-sorted portfolios (Panels C and D) for the early (1931:3–1963:2) and modern (1963:3–2011:4) subsamples. "Growth" denotes the lowest ratio of book equity to market equity (BE/ME), "Value," the highest BE/ME; "Small," the lowest market equity (ME); and "Large," the highest ME stocks. $\widehat{b}_{\Delta VAR}$ and $\widehat{b}_{r_M}$ are past return loadings on the weighted sum of changes in the vector autoregression (VAR) state variables, in which the weights are according to λ_V as estimated in Table 2, and on the market-return shock. "Diff" is the difference between the extreme cells. Bootstrapped standard errors [in brackets] are conditional on the estimated news series. Estimates are based on quarterly data using weighted least squares in which the weights are the same as those used to estimate the VAR.

Panel A: 25 ME– and BE/ME–sorted portfolios, early period (1931:3–1963:2)

	Growth	3	Value	Diff
$\widehat{\beta}_{CF}$				
Small	0.49	0.44	0.46	−0.04
	[0.13]	[0.11]	[0.10]	[0.05]
3	0.32	0.34	0.47	0.15
	[0.08]	[0.09]	[0.12]	[0.05]
Large	0.24	0.27	0.40	0.16
	[0.07]	[0.09]	[0.29]	[0.04]
Diff	−0.26	−0.17	−0.06	
	[0.07]	[0.04]	[0.03]	
$\widehat{\beta}_{DR}$				
Small	1.20	1.20	1.13	−0.07
	[0.15]	[0.17]	[0.17]	[0.07]
3	0.95	0.97	1.22	0.27
	[0.13]	[0.12]	[0.16]	[0.09]
Large	0.70	0.80	0.90	0.20
	[0.08]	[0.12]	[0.12]	[0.13]
Diff	−0.50	−0.40	−0.23	
	[0.14]	[0.16]	[0.08]	
$\widehat{\beta}_{V}$				
Small	−0.14	−0.15	−0.14	0.00
	[0.05]	[0.05]	[0.04]	[0.02]
3	−0.09	−0.09	−0.14	−0.05
	[0.03]	[0.03]	[0.04]	[0.02]
Large	−0.05	−0.09	−0.11	−0.07
	[0.02]	[0.04]	[0.03]	[0.03]
Diff	0.09	0.06	0.03	
	[0.04]	[0.02]	[0.02]	

Panel B: 25 ME– and BE/ME–sorted portfolios, modern period (1963:3–2011:4)

	Growth	3	Value	Diff
$\widehat{\beta}_{CF}$				
Small	0.23	0.26	0.28	0.05
	[0.06]	[0.05]	[0.05]	[0.04]
3	0.21	0.24	0.27	0.06
	[0.05]	[0.05]	[0.05]	[0.03]
Large	0.15	0.18	0.20	0.05
	[0.04]	[0.03]	[0.04]	[0.03]
Diff	−0.08	−0.08	−0.07	
	[0.04]	[0.03]	[0.03]	
$\widehat{\beta}_{DR}$				
Small	1.30	0.87	0.86	−0.44
	[0.11]	[0.07]	[0.09]	[0.08]
3	1.11	0.73	0.69	−0.42
	[0.08]	[0.06]	[0.07]	[0.08]
Large	0.82	0.60	0.64	−0.18
	[0.05]	[0.05]	[0.06]	[0.06]
Diff	−0.48	−0.26	−0.23	
	[0.10]	[0.06]	[0.08]	
$\widehat{\beta}_{V}$				
Small	0.13	0.05	0.01	−0.13
	[0.07]	[0.05]	[0.07]	[0.03]
3	0.14	0.05	0.04	−0.10
	[0.06]	[0.05]	[0.04]	[0.03]
Large	0.09	0.03	0.02	−0.08
	[0.05]	[0.04]	[0.04]	[0.02]
Diff	−0.04	−0.02	0.01	
	[0.03]	[0.02]	[0.03]	

(*continued on next page*)

Table 3 (*continued*)

Panel C: Six risk-sorted portfolios, early period (1931:3–1963:2)				
	Lo $\widehat{b}_{r_M}$	2	Hi $\widehat{b}_{r_M}$	Diff
$\widehat{\beta}_{CF}$				
Lo $\widehat{b}_{VAR}$	0.23	0.34	0.42	0.19
	[0.07]	[0.09]	[0.11]	[0.04]
Hi $\widehat{b}_{VAR}$	0.21	0.28	0.41	0.20
	[0.06]	[0.08]	[0.11]	[0.05]
Diff	−0.02	−0.05	−0.01	
	[0.02]	[0.03]	[0.02]	
$\widehat{\beta}_{DR}$				
Lo $\widehat{b}_{VAR}$	0.60	0.89	1.13	0.54
	[0.06]	[0.11]	[0.13]	[0.11]
Hi $\widehat{b}_{VAR}$	0.58	0.83	1.11	0.54
	[0.07]	[0.10]	[0.16]	[0.13]
Diff	−0.02	−0.06	−0.02	
	[0.04]	[0.08]	[0.06]	
$\widehat{\beta}_{V}$				
Lo $\widehat{b}_{VAR}$	−0.04	−0.07	−0.10	−0.06
	[0.02]	[0.03]	[0.04]	[0.02]
Hi $\widehat{b}_{VAR}$	−0.05	−0.07	−0.11	−0.06
	[0.02]	[0.03]	[0.04]	[0.03]
Diff	−0.01	0.00	−0.01	
	[0.02]	[0.02]	[0.02]	

Panel D: Six risk-sorted portfolios, modern period (1963:3–2011:4)				
	Lo $\widehat{b}_{r_M}$	2	Hi $\widehat{b}_{r_M}$	Diff
$\widehat{\beta}_{CF}$				
Lo $\widehat{b}_{VAR}$	0.20	0.20	0.26	0.06
	[0.04]	[0.04]	[0.06]	[0.04]
Hi $\widehat{b}_{VAR}$	0.17	0.21	0.21	0.05
	[0.03]	[0.04]	[0.06]	[0.05]
Diff	−0.04	0.01	−0.05	
	[0.03]	[0.02]	[0.02]	
$\widehat{\beta}_{DR}$				
Lo $\widehat{b}_{VAR}$	0.63	0.79	1.18	0.56
	[0.06]	[0.06]	[0.09]	[0.08]
Hi $\widehat{b}_{VAR}$	0.58	0.85	1.24	0.66
	[0.06]	[0.05]	[0.09]	[0.11]
Diff	−0.04	0.06	0.06	
	[0.09]	[0.06]	[0.05]	
Lo $\widehat{b}_{VAR}$	0.04	0.06	0.09	0.05
	[0.05]	[0.05]	[0.07]	[0.03]
Hi $\widehat{b}_{VAR}$	0.06	0.09	0.12	0.06
	[0.04]	[0.05]	[0.07]	[0.04]
Diff	0.02	0.03	0.03	
	[0.02]	[0.02]	[0.02]	

growth stocks relative to value stocks, growth stocks are relative hedges for two key aspects of the investment opportunity set. Consistent with CV (2004), growth stocks hedge news about future real stock returns. The novel finding of this paper is that growth stocks also hedge news about the variance of the market return.

One interesting aspect of these findings is the fact that the average β_V of the 25 size- and book-to-market portfolios changes sign from the early to the modern subperiod. Over the 1931–1963 period, the average β_V is -0.10, and over the 1964–2011 period this average becomes 0.06. Given the strong positive link between *PE* and volatility news documented in Panel D of Table 2, one should not be surprised that the market's β_V can be positive. Nevertheless, in the Online Appendix we study this change in sign more carefully. We show that the market's beta with realized volatility has remained negative in the modern period, highlighting the important distinction between realized and expected future volatility. We also show that the change in the sign of β_V is driven by a change in the correlation between the aggregate market return and the change in *DEFO*, our simple proxy for news about long-horizon variance.

5.2.2. Risk-sorted portfolios

Panels C and D of Table 3 show the estimated betas for the six risk-sorted portfolios over the 1931–1963 and post-1963 periods. In the pre-1963 sample period, high market-beta stocks have both higher cash-flow and higher discount-rate betas than low market-beta stocks. Low volatility-beta stocks have higher cash-flow betas and discount-rate betas than high volatility-beta stocks. High market-beta stocks also have lower volatility betas, but sorting stocks by their past volatility betas induces little spread in post-formation volatility betas. Putting these results together, in the 1931–1963 period, high market-beta stocks and low volatility-beta stocks were unambiguously riskier than low market-beta and high volatility-beta stocks.

In the post-1963 (modern) period, high market-beta stocks again have higher cash-flow and higher discount-rate betas than low market-beta stocks. However, high market-beta stocks now have higher volatility betas and are therefore safer in this dimension. This pattern perhaps is not surprising given our finding that the aggregate market portfolio itself has a positive volatility beta in the modern period. The important implication is that our three-beta model with priced volatility risk helps to explain the well-known result that stocks with high past market betas have offered relatively little extra return in the past 50 years (Fama and French, 1992; Frazzini and Pedersen, 2013).

In the modern period, sorts on volatility beta generate an economically and statistically significant spread in post-formation volatility beta. These high volatility-beta portfolios also tend to have higher discount-rate betas and lower cash-flow betas, though the patterns are not uniform.

We also examine test assets that are formed based on both characteristics and risk estimates. The Online Appendix reports the estimated betas for the 18 BE/ME-*ivol* $-$ $\widehat{\beta}_{\Delta VAR}$-sorted portfolios in both the early and modern sample periods. In the early period, firms with higher *ivol* have lower post-formation volatility betas regardless of their book-to-market ratio. Consistent with this finding, higher *ivol* stocks have higher average returns. In the modern period, however, among stocks with low BE/ME, firms with higher *ivol* have higher post-formation volatility betas and lower average return. These patterns reverse among stocks with high BE/ME.

We argue that these differences make economic sense. High idiosyncratic volatility increases the value of growth options, which is an important effect for growing firms with flexible real investment opportunities, but much less so for stable, mature firms. Valuable growth options in turn imply high betas with aggregate volatility shocks. Hence, high idiosyncratic volatility naturally raises the volatility beta for growth stocks more than for value stocks. This effect is stronger in the modern sample in which growing firms with flexible investment opportunities are more prevalent.

Taken together, the findings from the characteristic- and risk-sorted test assets suggest that volatility betas vary with multiple stock characteristics and that techniques taking this into account can be more effective in generating a spread in post-formation volatility beta.

5.3. Model estimation

We now turn to pricing the cross section of excess returns on our test assets. We estimate our model's single parameter via GMM, using the moment condition Eq. (15). For ease of exposition, we report our results in terms of the expected return-beta representation from Eq. (16), rescaled by the variance of market return innovations as in Section 5.2:

$$\overline{R}_i - \overline{R}_j = g_1 \widehat{\beta}_{i,CF_M} + g_2 \widehat{\beta}_{i,DR_M} + g_3 \widehat{\beta}_{i,V_M} + e_i, \tag{24}$$

where bars denote time series means and betas are measured using returns relative to the reference asset. Recall that we use the aggregate equity market as our reference asset but include the T-bill return as a test asset, so that our model not only prices cross-sectional variation in average returns, but also prices the average difference between stocks and bills.

We evaluate the performance of five asset pricing models, all estimated via GMM: (1) the traditional CAPM that restricts cash-flow and discount-rate betas to have the same price of risk and sets the price of variance risk to zero, (2) the two-beta intertemporal asset pricing model of CV (2004) that restricts the price of discount-rate risk to equal the variance of the market return and again sets the price of variance risk to zero, (3) our three-beta intertemporal asset pricing model that restricts the price of discount-rate risk to equal the variance of the market return and constrains the prices of cash-flow and variance risk to be related by Eq. (10), with $\rho = 0.95$ per year, (4) a partially constrained three-beta model that restricts the price of discount-rate risk to equal the variance of the market return but freely estimates the other two risk prices (effectively decoupling γ and ω), and (5) an unrestricted three-beta model that allows free risk prices for cash-flow, discount-rate, and volatility betas.

5.3.1. Model estimates

Table 4 reports the results of pricing tests for both the early sample period 1931–1963 (Panel A) and the modern sample period 1963–2011 (Panel B). In each case, we price the complete set of test assets described in Section 5.1. The Online Appendix reports the results of tests that price the 25 size- and book-to-market–sorted portfolios in isolation. Table 4 has five columns, one for each of our asset pricing models. The first six rows of each panel are divided into three sets of two rows. The first set of two rows corresponds to the premium on cash-flow beta, the second set to the premium on discount-rate beta, and the third set to the premium on volatility beta. Within each set, the first row reports the point estimate in fractions per quarter, and the second row reports the corresponding standard error. Below the premia estimates, we report the R^2 statistic for a cross-sectional regression of average market-adjusted returns on our test assets onto the fitted values from the model as well as the J statistic. In the next two rows of each panel, we report the implied risk-aversion coefficient, γ, which can be recovered as g_1/g_2, as well as the sensitivity of news about risk to news about market variance, ω, which can be recovered as $-2g_3/g_2$. The five final rows in each panel report the cross-sectional R^2 statistics for various subsets of the test assets.

Table 4, Panel A, shows that, in the early subperiod, all models do a relatively good job pricing these one hundred test assets. The cross-sectional R^2 statistic is 74% for the CAPM, 78% for the two-beta ICAPM, and 79% for our three-beta ICAPM. Consistent with the claim that the three-beta model does a good job describing the cross section, the constrained and the unrestricted factor model barely im-

Table 4
Asset pricing tests.
The table reports generalized methods of moments estimates of the capital asset pricing model (CAPM), the two-beta intertemporal CAPM (ICAPM), the three-beta volatility ICAPM, a factor model in which only the $\widehat{\beta}_{DR}$ premium is restricted, and an unrestricted factor model for the early (Panel A: 1931:3–1963:2) and modern (Panel B: 1963:3–2011:4) subsamples. The test assets are 25 market equity (ME)- and ratio of book equity to market equity (BE/ME)-sorted portfolios and the T-bill, six risk-sorted portfolios, 18 characteristic- and risk-sorted assets, and managed versions of these portfolios, scaled by *EVAR*, and the reference asset is the market portfolio. The 5% critical value for the test of overidentifying restrictions is 121.0 in Columns 1, 2, and 3; 119.9 in Column 4; and 118.8 in Column 5. N/A = not applicable for the model in question.

Parameter	CAPM	Two-beta ICAPM	Three-beta ICAPM	Constrained	Unrestricted
	(1)	(2)	(3)	(4)	(5)
Panel A: early period					
$\widehat{\beta}_{CF}$ premium (g_1)	0.037	0.105	0.081	0.058	0.101
(Standard error)	(0.016)	(0.071)	(0.037)	(0.052)	(0.067)
$\widehat{\beta}_{DR}$ premium (g_2)	0.037	0.016	0.016	0.016	−0.016
(Standard error)	(0.016)	(0)	(0)	(0)	(0.017)
$\widehat{\beta}_{VAR}$ premium (g_3)			−0.049	−0.094	−0.197
(Standard error)			(0.068)	(0.126)	(0.142)
$\widehat{R}^2$	74%	78%	79%	79%	81%
J statistic	735.9	844.6	824.7	811.1	849.4
Implied γ	2.4	6.6	5.1	N/A	N/A
Implied ω	N/A	N/A	6.2	N/A	N/A
$\widehat{R}^2$: 26 unscaled characteristics	64%	66%	67%	68%	69%
$\widehat{R}^2$: six unscaled risk	57%	35%	53%	67%	73%
$\widehat{R}^2$: 18 unscaled characteristics and risk	67%	73%	75%	75%	83%
$\widehat{R}^2$: 50 unscaled	66%	68%	70%	71%	74%
$\widehat{R}^2$: 50 scaled	67%	72%	73%	74%	77%
Panel B: Modern period					
$\widehat{\beta}_{CF}$ premium (g_1)	0.014	0.118	0.055	0.099	0.104
(Standard error)	(0.010)	(0.056)	(0.000)	(0.040)	(0.030)
$\widehat{\beta}_{DR}$ premium (g_2)	0.014	0.008	0.008	0.008	0.004
(Standard error)	(0.010)	(0)	(0)	(0)	(0.014)
$\widehat{\beta}_{VAR}$ premium (g_3)			−0.096	−0.120	−0.116
(Standard error)			(0.035)	(0.034)	(0.041)
$\widehat{R}^2$	−20%	25%	60%	71%	72%
J statistic	499.2	364.7	495.3	383.8	342.0
Implied γ	1.9	15.2	7.2	N/A	N/A
Implied ω	N/A	N/A	24.9	N/A	N/A
$\widehat{R}^2$: 26 unscaled characteristic portfolios	−51%	45%	48%	74%	73%
$\widehat{R}^2$: six unscaled risk portfolios	−10%	23%	49%	71%	67%
$\widehat{R}^2$: 18 unscaled characteristic and risk portfolios	−27%	26%	62%	71%	75%
$\widehat{R}^2$: 50 unscaled portfolios	−31%	36%	57%	73%	75%
$\widehat{R}^2$: 50 scaled portfolios	−16%	17%	62%	69%	69%

prove pricing relative to the three-beta ICAPM in Panel A. Despite this apparent success, all models are rejected based on the standard *J* test. This perhaps is not surprising, given that even the empirical three-factor model of Fama and French (1993) is rejected by this test when faced with the 25 size- and book-to-market–sorted portfolios.

In stark contrast, Panel B shows that, in the modern subperiod, the CAPM fails to price not only the characteristic-sorted test assets already considered in previous work, but also risk-sorted and variance-scaled portfolios. The cross-sectional R^2 of the CAPM is negative at −20%. The two-beta ICAPM of CV (2004) does a better job describing average returns in the modern subperiod, delivering an R^2 of 25%, but it struggles to price the risk-sorted and variance-scaled test assets and requires a much larger coefficient of risk aversion in the modern subperiod than in the early subperiod.

In the modern period, the three-beta ICAPM outperforms both the CAPM and the two-beta ICAPM, delivering an overall R^2 of 60%. The model also does a good job explaining all the subsets of test assets that we consider, including the risk-sorted and variance-scaled test assets. Moreover, the three-beta estimate of risk aversion is relatively stable across subperiods. This improvement is driven by the addition of volatility risk to the model. Our estimate of the volatility premium is both economically and statistically significant. The premium for one unit of volatility beta is approximately −38% per year and 2.76 standard deviations from zero.

Further support for our three-beta ICAPM can be found in the last two columns. Relaxing the link between γ and ω (but continuing to restrict the premium for discount-rate beta) only improves the fit somewhat (from 60% to 71%). The γ and ω of the partially constrained model are 12.2 and 31.0, respectively, which are not dramatically different from the estimated parameters of the fully constrained version of the model. Furthermore, a completely unrestricted three-beta model has an R^2 (72%) that is very close to that of the partially constrained implementation. Finally, the premium for variance beta is relatively stable and always statistically significant across all three versions of our three-beta model (ICAPM, partially constrained, and unrestricted).

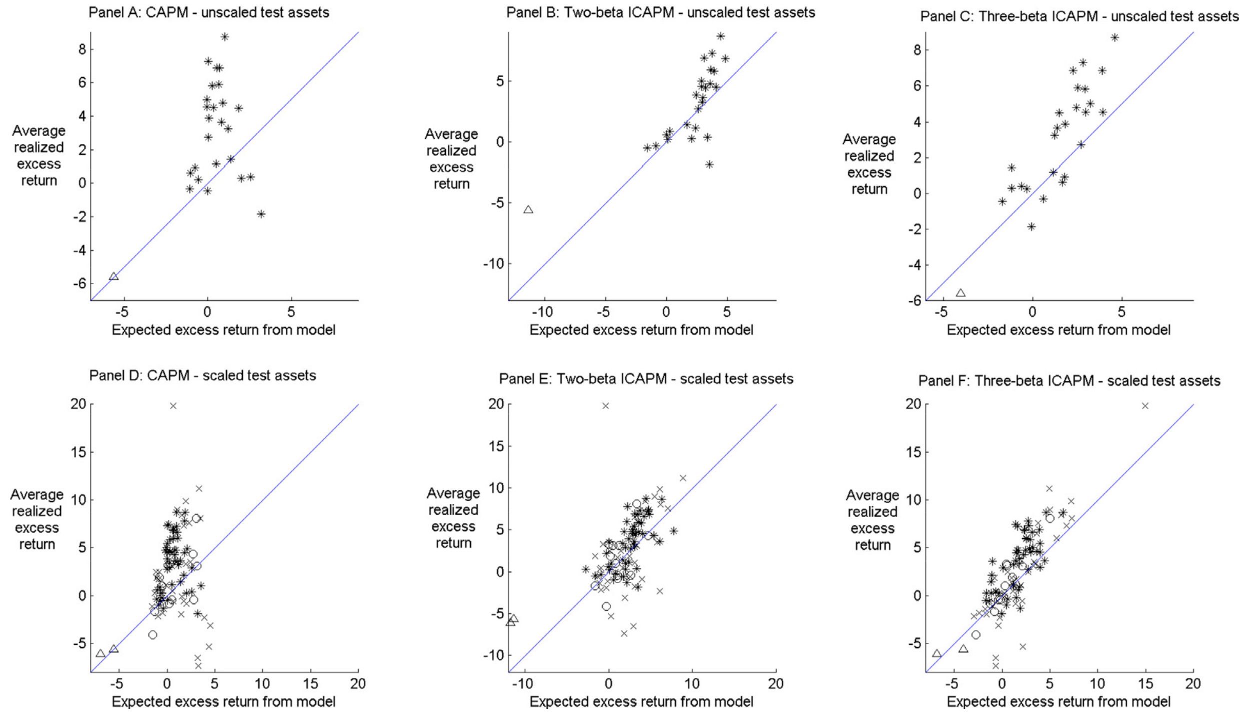

Fig. 4. Pricing tests: Each diagram plots the sample against predicted average excess returns. Test assets in Panels A–C are the 25 market equity (ME)- and ratio of book equity to market equity (BE/ME)-sorted portfolios (asterisks), plus the bill return (triangle); and in Panels D–F, both unscaled and scaled by *EVAR* versions of the 25 ME- and BE/ME-sorted portfolios (asterisks), six risk-sorted portfolios (circles), 18 characteristic- and risk-sorted portfolios (crosses), and bill return (triangles). Predicted values are from Table 4 for 1963:3–2011:4. The models tested are the capital asset pricing model (CAPM), the two-beta intertemporal CAPM, and the three-beta ICAPM.

Fig. 4 provides a visual summary of the modern-period results reported in Table 4, Panel B. Each panel in the figure plots average realized excess returns against average predicted excess returns from one of the asset pricing models under consideration. A well-specified model should deliver points that lie along the 45-degree line when realized returns are measured over a long enough sample period.

In Panels A-C of Fig. 4, we examine how these models price the original 25 characteristic-sorted portfolios, which are plotted as stars, along with the Treasury bill, plotted as a triangle. The CAPM is plotted in Panel A, the two-beta ICAPM, in Panel B, and the three-beta ICAPM, in Panel C. The poor performance of the CAPM in this sample period and the increase in explanatory power provided by the two-beta ICAPM and particularly the three-beta ICAPM are immediately apparent. The two-beta ICAPM has particular difficulty with the Treasury bill, predicting far too low an excess return relative to the aggregate stock market or, equivalently, far too high an equity premium. Panels D-F of Fig. 4 provide a summary of the modern-period results with the full set of test assets. A visually striking improvement in fit is evident moving from the CAPM to the two-beta ICAPM and then to the three-beta ICAPM.

5.3.2. Implications for the history of marginal utility

As a way to understand the economics behind the ICAPM, and as a further check on the reasonableness of our model, we consider what the model implies for the history of our investor's marginal utility. Fig. 5 plots the time series of the combined shock $\gamma N_{CF} - N_{DR} - \frac{1}{2}\omega N_V$, normalized and then smoothed for graphical purposes as in Fig. 2, based on our estimate of the three-beta model using characteristic-sorted test assets in the modern period (Table 4, Panel B). The smoothed shock has correlation 0.77 with equivalently smoothed N_{CF}, 0.02 with smoothed $-N_{DR}$, and -0.80 with smoothed N_V. Fig. 5 also plots the corresponding smoothed shock series for the CAPM ($N_{CF} - N_{DR}$) and for the two-beta ICAPM ($\gamma N_{CF} - N_{DR}$). The two-beta model shifts the history of good and bad times relative to the CAPM, as emphasized by CGP (2013). The model with stochastic volatility further accentuates that periods with high market volatility, such as the 1930s and the late 2000s, are particularly hard times for long-term investors. Assets that do well in such hard times, for example, growth stocks, are valuable hedges that should have low average returns.

6. An ICAPM perspective on asset pricing anomalies

In this section, we use our ICAPM model to reassess a wide variety of anomalies that have been discussed in the asset pricing literature. We begin with equity anomalies and then consider some anomalous patterns from outside the equity market.

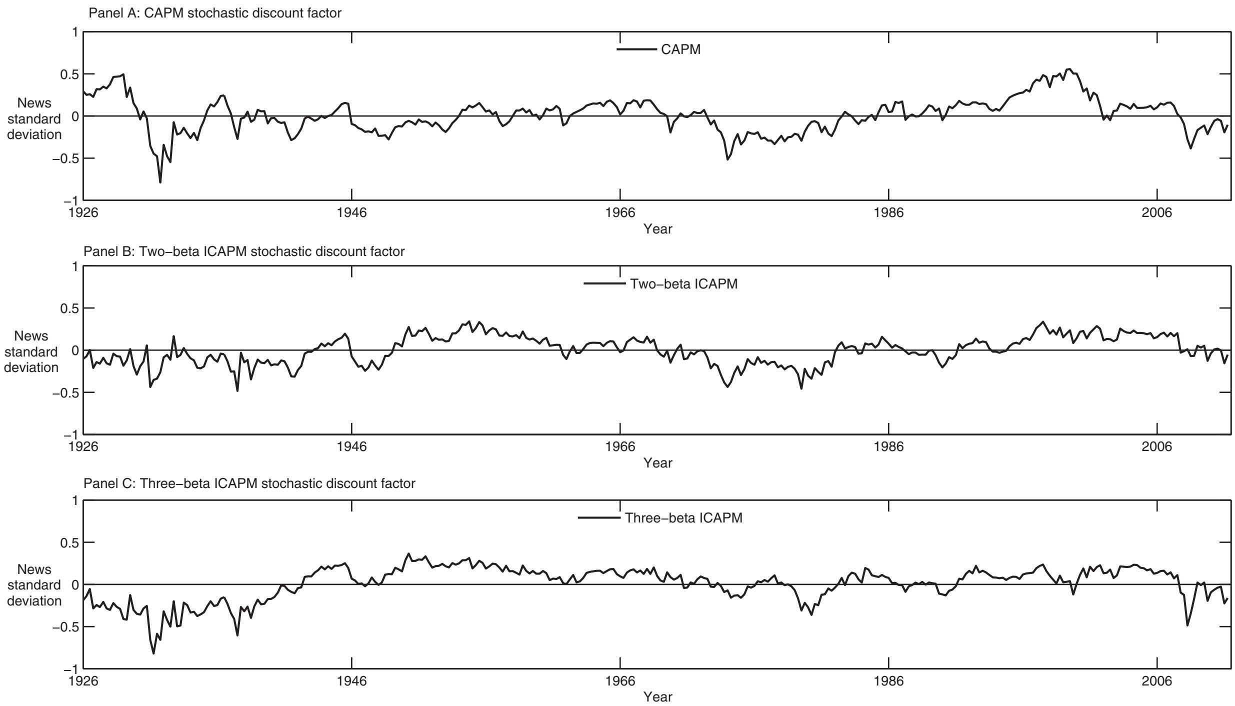

Fig. 5. Stochastic discount factor shocks across models: This figure plots the time series of the smoothed combined shock for the capital asset pricing model (CAPM) ($N_{CF} - N_{DR}$), the two-beta intertemporal CAPM ($\gamma N_{CF} - N_{DR}$), and the three-beta ICAPM that includes stochastic volatility ($\gamma N_{CF} - N_{DR} - \frac{1}{2}\omega N_V$) estimated in Table 4, Panel B, for the sample period 1963:3-2011:4. For each model, the shock is normalized by its standard deviation and then smoothed with a trailing exponentially weighted moving average. The decay parameter is set to 0.08 per quarter, and the smoothed normalized shock series is generated as $MA_t(SDF) = 0.08SDF_t + (1 - 0.08)MA_{t-1}(SDF)$. This decay parameter implies a half-life of approximately two years.

6.1. Equity anomalies

Table 5 analyzes a number of well-known equity anomalies using data taken from Professor Kenneth French's website.[16] The sample period is 1963:3–2011:4. The anomaly portfolios are the market (*RMRF*), size (*SMB*), and value (*HML*) equity factors of Fama and French (1993), the profitability (*RMW*) and investment (*CMA*) factors added in Fama and French (2016), the momentum (*UMD*) factor of Carhart (1997), short-term reversal (*STR*) and long-term reversal (*LTR*) factors, and zero-cost portfolios formed from value-weighted quintiles sorted on beta (*BETA*), accruals (*ACC*), net issuance (*NI*), and idiosyncratic volatility (*IVOL*). We also consider a dynamic portfolio that varies its exposure to the equity premium based on c/PE_t, where c is chosen so that the resulting managed portfolio has the same unconditional volatility as *RMRF*. We refer to this portfolio as *MANRMRF*.

For each of these portfolios, Panel A reports the mean excess return in the first column and the standard deviation of return in the second column. Columns 3–5 report the portfolios' betas with our estimates of discount-rate news, cash-flow news, and variance news. These are used in Columns 6–9 to construct the components of fitted excess returns based on discount-rate news (λ_{DR}), cash-flow news in the two-beta ICAPM (λ_{CF}^{2-BETA}), cash-flow news in the three-beta ICAPM (λ_{CF}^{3-BETA}), and variance news in the three-beta ICAPM (λ_V). These fitted excess returns use the parameter estimates of the two-beta and three-beta models reported in Table 4, Panel B. We do not reestimate any parameters and in this sense the evaluation of equity anomalies is out of sample.

Columns 10–12 of Panel A report the alphas of the anomalies (their sample average excess returns less their predicted excess returns) calculated using the CAPM, the two-beta ICAPM, and the three-beta ICAPM. All the portfolios, with the obvious exception of *RMRF*, have been chosen to have positive CAPM alphas. The ability of the ICAPM to explain asset pricing anomalies can be measured by the reduction in magnitude of ICAPM alphas relative to CAPM alphas. To summarize model performance, Panel B reports average absolute alphas across all anomaly portfolios, the three Fama and French (1993) portfolios, and the five Fama and French (2016) portfolios. These averages are calculated both for raw alphas and after dividing each anomaly's alpha by the standard deviation of its return.

Table 5 shows that volatility risk exposure is helpful in explaining many of the equity anomalies that have been discussed in the recent asset pricing literature. Most of the anomaly portfolios have negative variance betas, which make them riskier and help to explain their positive excess returns. Exceptions to this statement include the excess return on the market over a Treasury bill *RMRF* and the managed excess return *MANRMRF* (as we find the market to be a volatility hedge in the modern sub-

[16] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

Table 5
Pricing popular equity strategies.
The table decomposes the average quarterly returns on well-known equity strategies using the capital asset pricing model (CAPM), the two-beta intertemporal CAPM (ICAPM), and our three-beta ICAPM. We estimate α_{CAPM} using a standard time series regression. We estimate α^{2-BETA}_{ICAPM} and α^{3-BETA}_{ICAPM} using the corresponding estimates of γ from Table 4, Panel B. The sample covers the 1963:3–2011:4 time period during which the market variance is 0.0077. The strategies are the market (RMRF), size (SMB), value (HML), profitability (RMW), investment (CMA), momentum (UMD), short-term reversal (STR), and long-term reversal (LTR) factors as well as zero-cost portfolios formed from value-weight quintiles sorted on beta (BETA), accruals (ACC), net issuance (NI), or idiosyncratic volatility (IVOL). We also consider a dynamic portfolio that varies its exposure to the equity premium based on $\frac{c}{PE_t}$, where c is chosen so that the resulting managed portfolio has the same unconditional volatility as *RMRF*. We refer to this portfolio as *MANRMRF*. All return data are from Ken French's website, available at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/. We report the average absolute model αs for various subsets of the strategies, considering not only the raw strategies but also when the strategies are rescaled to have the same volatility as *RMRF*. As part of the comparison, we also calculate model αs using the constrained and unrestricted models of Table 4, Panel B, as well as the three- and five-factor models of Fama and French.

Panel A: Decomposition of average quarterly returns

Strategy	μ(%) (1)	σ(%) (2)	$\widehat{\beta}_{DR}$ (3)	$\widehat{\beta}_{CF}$ (4)	$\widehat{\beta}_{V}$ (5)	λ_{DR} (%) (6)	λ^{2-BETA}_{CF} (%) (7)	λ^{3-BETA}_{CF} (%) (8)	λ_{V} (%) (9)	α_{CAPM} (%) (10)	α^{2-BETA}_{ICAPM} (%) (11)	α^{3-BETA}_{ICAPM} (%) (12)
RMRF	1.39	8.69	0.78	0.19	0.07	0.60	2.25	1.06	−0.70	0	−1.45	0.44
SMB	0.78	5.65	0.22	0.06	0.02	0.17	0.67	0.32	−0.17	0.35	−0.07	0.45
HML	1.18	5.92	−0.26	0.05	−0.10	−0.20	0.55	0.26	0.94	1.50	0.83	0.18
RMW	0.83	4.17	−0.09	−0.01	0.01	−0.07	−0.14	−0.07	−0.10	0.99	1.04	1.06
CMA	1.02	4.21	−0.21	0.02	−0.05	−0.16	0.22	0.10	0.47	1.30	0.96	0.61
UMD	2.18	7.78	−0.14	−0.03	0.03	−0.11	−0.35	−0.16	−0.26	2.46	2.64	2.71
BETA	−0.20	10.90	−0.74	−0.08	−0.05	−0.57	−0.91	−0.43	0.50	1.01	1.28	0.30
STR	1.58	5.66	0.15	0.05	−0.01	0.12	0.55	0.26	0.07	1.28	0.91	1.14
LTR	0.92	5.27	−0.09	0.05	−0.05	−0.07	0.56	0.26	0.47	0.97	0.43	0.26
ACC	1.14	4.29	−0.08	−0.03	−0.02	−0.06	−0.34	−0.16	0.21	1.29	1.54	1.15
NI	1.19	5.59	−0.21	−0.03	−0.02	−0.16	−0.33	−0.16	0.21	1.57	1.68	1.30
IVOL	1.02	11.61	−0.76	−0.07	−0.05	−0.58	−0.87	−0.41	0.52	2.26	2.47	1.50
MANRMRF	1.48	8.69	0.76	0.20	0.08	0.58	2.29	1.08	−0.74	0.10	−1.39	0.56

Panel B: Average absolute alpha

Strategy	α_{CAPM} (%)	α^{2-BETA}_{ICAPM} (%)	α^{3-BETA}_{ICAPM} (%)	α^{3-BETA}_{Constr} (%)	$\alpha^{3-BETA}_{Unrestr}$ (%)	α_{3FF}(%)	α_{5FF} (%)
All, not scaled	1.16	1.28	0.90	0.85	0.80	0.90	0.57
All, scaled	1.65	1.69	1.26	1.18	1.13	1.23	0.80
Three-factor model, not scaled	0.62	0.78	0.36	0.25	0.23	0	0
Three-factor model, scaled	0.91	0.93	0.47	0.33	0.34	0	0%
Five-factor model, not scaled	0.83	0.87	0.55	0.46	0.43	0.32	0%
Five-factor model, scaled	1.50	1.39	0.98	0.84	0.81	0.67	0%

period) and the returns on small size *SMB*, profitability *RMW*, and momentum *UMD*. The three-beta ICAPM is particularly good at explaining the high return on value *HML*, which perhaps is not surprising because we estimate the model using size- and value-sorted equity portfolios. But it also makes considerable progress at explaining the returns to low-investment firms *CMA*, low-beta stocks *BETA*, long-term reversal *LTR*, and low idiosyncratic volatility *IVOL*.

Across all the anomalies in the table, the average absolute alpha is 1.16% for the CAPM, slightly higher at 1.28% for the two-beta ICAPM, but lower at 0.90% for the three-beta ICAPM. Looking only at Fama and French (1993) anomalies, the three-beta model reduces the average absolute alpha from the CAPM's 0.62% to 0.36%. Looking only at Fama and French (2016) anomalies the average absolute alpha falls from 0.83% to 0.55%. In both these subsets, the two-beta ICAPM actually performs worse than the CAPM. Results are similar when anomaly returns are scaled by standard deviation.

To what extent is our progress substantial? One reasonable way to gauge these results is by comparing the pricing improvement (relative to the CAPM) of our model with unrestricted models of the risk-return trade-off. Panel B of Table 5 provides exactly those comparisons. For example, one such possible benchmark is the unrestricted three-beta version of our model, with the factors N_{CF}, $-N_{DR}$, and N_V. Using only a single free parameter, our three-beta ICAPM provides 72% of the pricing improvement that an unrestricted multifactor model does. Other reasonable benchmarks studied in the table are the three- and five-factor models of Fama and French (1993, 2016). Relative to those models, our three-beta ICAPM provides 100% and 44% of the respective pricing improvement. Of course, that class of models is built from portfolios directly sorted on several of the anomalies studied in Table 5, which makes our pricing improvement even more impressive.

6.2. Non-equity anomalies

Table 6 considers several sets of non-equity test assets, each of which is measured from a different start date until the end of our sample period in 2011:4. First, we consider $HY - IG$, the risky bond factor of Fama and French (1993), which we measure from 1983:3 using the return on the Barclays Capital High Yield Bond Index (*HYRET*) less the return on Barclays Capital Investment Grade Bond Index (*IGRET*). Second, we study the cross section of currency portfolios (*CARRY*), starting in 1984:1, with developed-country currencies dynamically allocated

Table 6
Pricing popular non-equity strategies.
The table decomposes the average quarterly returns on well-known non-equity strategies using the capital asset pricing model (CAPM), the two-beta intertemporal CAPM, and our three-beta ICAPM. We estimate α_{CAPM} using a standard time series regression. We estimate α^{2-BETA}_{ICAPM} and α^{3-BETA}_{ICAPM} using the corresponding estimates of γ from Table 4, Panel B. The strategies are a risky bond factor (HY-IG) that buys high yield bonds and shorts investment-grade bonds, a carry factor (CARRY) from the cross-section of developed-country currencies, a short position in an Standard & Poor's (S&P) 100 index straddle (STRADDLE), and a term bet on S&P 500 synthetic variance forward (VIXF2-VIXF0). The sample periods and market variance (in parentheses) corresponding to these four strategies are 1983:3–2011:4 (0.0077), 1984:1–2011:4 (0.0078), 1986:1–2011:4 (0.0080), and 1998:1–2011:4 (0.0101). The text provides more details on the source of each of these four non-equity strategies. We report the average absolute model αs for various subsets of the strategies, considering not only the raw strategies but also when the strategies are rescaled to have the same volatility as *RMRF*. As part of the comparison, we also calculate model αs using the constrained and unrestricted models of Table 4, Panel B, as well as the three- and five-factor models of Fama and French.

Panel A: Decomposition of average quarterly returns

Strategy	μ (%)	σ (%)	$\widehat{\beta}_{DR}$	$\widehat{\beta}_{CF}$	$\widehat{\beta}_{V}$	λ_{DR} (%)	λ^{2-BETA}_{CF} (%)	λ^{3-BETA}_{CF} (%)	λ_{V} (%)	α_{CAPM} (%)	α^{2-BETA}_{ICAPM} (%)	α^{3-BETA}_{ICAPM} (%)
HY-IG	0.23	4.47	0.25	0.01	−0.06	0.19	0.15	0.07	0.61	−0.27	−0.12	−0.65
CARRY	1.48	5.37	0.19	0.01	−0.07	0.15	0.12	0.06	0.72	1.11	1.20	0.55
STRADDLE	21.66	47.10	1.90	0.18	−0.29	1.53	2.19	1.03	2.89	17.71	17.94	16.21
VIXF2-VIXF0	26.84	48.41	2.74	0.24	−0.25	2.77	3.66	1.72	3.20	24.56	20.40	19.14

Panel B: Average absolute alpha

Strategy	α_{CAPM} (%)	α^{2-BETA}_{ICAPM} (%)	α^{3-BETA}_{ICAPM} (%)	α^{3-BETA}_{Constr} (%)	$\alpha^{3-BETA}_{Unrestr}$ (%)	α_{3FF} (%)	α_{5FF} (%)
All, not scaled	10.91	9.92	9.14	8.69	9.30	10.73	12.01
All, scaled	2.50	2.29	2.15	2.08	2.18	2.43	2.53

to portfolios based on their interest rates as in Lustig et al. (2011).[17] Third, we use the S&P 100 index straddle returns (*STRADDLE*) studied by Coval and Shumway (2001), starting in 1986:1.[18]

Finally, from the S&P 500 options market, we generate quarterly returns on three synthetic variance forward contracts starting in 1998:3. We construct these returns as in Dew-Becker et al. (2016). We construct a panel of implied variance swap prices using option data from OptionMetrics, for maturities n ranging from one quarter to three quarters ahead: $VIX^2_{n,t}$. Under the assumption that returns follow a diffusion, we have $VIX^2_{n,t} = E^Q_t[\int_t^{t+n}\sigma^2_s ds]$. We compute $VIX^2_{n,t}$ using the same methodology used by the Chicago Board Options Exchange to construct the 30-day Volatility Index (VIX), applying it to maturities up to three quarters. We then compute synthetic variance forward prices as $F_{n,t} = VIX^2_{n,t} - VIX^2_{n-1,t}$. These forward allow us to isolate claims to variance at a specific horizon n (focusing on the variance realized between $n-1$ and n). The quarterly returns to these forward are computed as $R_{n,t} = \frac{F_{n-1,t}}{F_{n,t-1}} - 1$, where $F_{0,t} = RVAR_t$. Dew-Becker et al. (2016) find a large difference in average returns for these forward across maturities. Accordingly, we construct the anomaly portfolio as a long-short portfolio that sells short-maturity forward and buys long-maturity forward (yielding strongly positive average returns).

All these anomaly portfolios have been normalized to have positive excess returns, and they all have negative variance betas so their exposure to variance risk does contribute to an explanation of their positive returns. However, in the case of $HY-IG$, the three-beta model overshoots and predicts a higher average return than has been realized in the data. In the case of *CARRY*, the three-beta model cuts the CAPM alpha roughly in half. In the two options anomalies, *STRADDLE* and $VIXF2 - VIXF0$, the three-beta model reduces the CAPM alpha slightly but the high returns to these anomalies remain puzzling even after taking account of their long-run volatility risk exposures.

Though our three-beta ICAPM is far from perfect in absolute terms, our model fares relatively well compared with unrestricted asset pricing models. For example, the unrestricted version of our model has slightly higher average absolute pricing errors. Perhaps even more impressively, our economically motivated ICAPM significantly outperforms both the three- and five-factor versions of the empirical models of Fama and French.

These findings relate to the literature on the pricing of volatility risk in derivative markets (Coval and Shumway, 2001; Ait-Sahalia et al., 2015). Dew-Becker et al. (2016) study the market for variance swaps with different maturities and show that market risk premia associated with short-term variance shocks are highly negative, whereas risk premia for news shocks about future variance are close to zero. These results present a challenge to mod-

[17] We thank Nick Roussanov for sharing these data.

[18] The series we study includes only those straddle positions in which the difference between the options' strike price and the underlying price is between zero and five. We thank Josh Coval and Tyler Shumway for providing their updated data series to us.

els in which investors have strong intertemporal hedging motives, including our model and the long-run risk model of BKSY (2014). It perhaps is not surprising that the intertemporal model of this paper, which is based on the first-order conditions of a long-term equity investor, works better for equity anomalies than for anomalies in derivatives markets, which are harder to access for this type of investor.

7. Alternative specifications and robustness

In this section, we compare our model with some alternatives that have recently been explored in the literature. We also briefly discuss the robustness of our results to alternative choices in the empirical implementation.

7.1. Comparison with the BKSY (2014) model

In this subsection, we explore the main differences between our paper and BKSY (2014), regarding both modeling assumptions and empirical implementation.

A first difference lies in the modeling of the volatility process itself. In our paper, we model volatility as a heteroskedastic process. In their main results, BKSY employ a homoskedastic volatility process. A disadvantage of BKSY's specification is that the volatility process becomes negative more frequently than in the case of a heteroskedastic process, in which the volatility of innovations to volatility shrinks as volatility gets close to zero. In the Online Appendix we explore this difference formally, using simulations to compare the frequency with which the heteroskedastic and homoskedastic models become negative, showing a clear advantage in favor of the heteroskedastic process. If one adjusts the volatility process upward to zero whenever it would otherwise go negative, the cumulative adjustment required quickly decreases to zero for the heteroskedastic process as the sampling frequency increases. It does not for the homoskedastic process. In our simulations, the ratio of the adjustment needed in the homoskedastic case relative to the one needed in the heteroskedastic case is 6 at the quarterly frequency, 17 at the monthly frequency, and over 200 at the daily frequency.

BKSY's assumption of homoskedastic volatility has important consequences for their asset pricing analysis. In the Online Appendix, we show that if the volatility process is homoskedastic, the SDF can be expressed as a function of variance news N_V only under special conditions not explicitly stated by BKSY: that the N_V shock depends only on innovations to state variables that are themselves homoskedastic, and that N_{CF} and N_V are uncorrelated.[19] In our empirical analysis, we estimate the correlation between N_{CF} and N_V to be −0.12. We also explore a range of other specifications for the VAR and find that this correlation is often below −0.5 and, in some cases, as low as −0.78. In fact, when we emulate BKSY's VAR specification, we obtain a strongly negative correlation of −0.71. This result should not be surprising: the literature on the leverage effect (Black, 1976; Christie, 1982) has long shown that news about low cash flows is associated with news about higher future volatility. Overall, the empirical analysis provides strong evidence that assuming a zero correlation between N_{CF} and N_V, as BKSY implicitly do, is counterfactual across a range of specifications.

In a robustness exercise in their Sections II.E and III.D, BKSY (2014) entertain a heteroskedastic process similar to ours, in which a single variable σ_t^2 drives the conditional variance of all variables in the VAR. In this specification, no theoretical constraints are placed on the correlation between N_{CF} and N_V. However, as discussed in subsection 3.2.1, another constraint appears in models with heteroskedastic volatility, that is, the value function of the investor ceases to exist once risk aversion becomes sufficiently high. The most visible symptom of the existence issue is that the function that links ω (the price of risk of N_V) to risk aversion γ is not defined in this region. The condition for existence of a solution is a nonlinear function of the structural parameters of the model and the time series properties of the state variables. BKSY ignore the existence constraint by linearizing the function $\omega(\gamma)$ around $\gamma = 0$.[20] This approach has two problems. First, the empirical estimates of the model parameters may erroneously imply a model solution that lies in the nonexistence region. Second, even when the model is in a region of the parameter space where a solution would exist, BKSY's solution is based on an approximation whose accuracy is not clear and not explored in the paper.

In addition to these different modeling assumptions, BKSY (2014) differs from our paper in the empirical implementation. This leads to several important differences in the findings. First, we find that variance risk premia make an important contribution to explaining the cross section of equity returns, while they contribute only minimally in BKSY. Second, we find that a value-minus-growth bet has a negative beta with volatility news, while BKSY find that it has a positive volatility beta. Third, in the modern period, we estimate the aggregate stock market to have a positive volatility beta, while BKSY estimate a negative volatility beta.

To better understand the source of the differences in empirical results, the Online Appendix explores the properties of the news terms using different VAR specifications including our baseline specification, BKSY's baseline (for the part of their analysis expressed in terms of returns instead of consumption, so directly comparable to ours), and various combinations of those. We focus on three main differences in the empirical approach: (1) the estimation of a VAR at yearly versus quarterly frequencies, (2) the methodology used to construct realized variance since we construct realized variance using sum of squared daily returns, whereas BKSY use sums of squared monthly returns that ignore the information in higher-frequency data and result in a noisier estimator of realized variance, and (3) the

[19] Other knife-edge cases with a solution can exist even when N_{CF} and N_V are correlated, but they entail even more extreme assumptions, for example, N_V not loading at all on volatility innovations, or the set of news terms not depending at all on any heteroskedastic state variable. The Online Appendix provides details.

[20] In the first draft of our paper, we also used this inappropriate linearization.

use of different state variables, and particularly the value spread, that we show to be important for our results and that is not included in BKSY (2014). This analysis shows that both using high-frequency data to compute *RVAR* and including the value spread are important drivers of the differences between our results and those of BKSY.[21]

With regard to the difference in the estimated volatility beta of a value-minus-growth portfolio, we note that our negative volatility beta estimate is more consistent with models in which growth firms hold options that become more valuable when volatility increases (Berk et al., 1999; McQuade, 2012; Dou, 2016). Empirically, our negative volatility beta estimate is consistent with the underperformance of value stocks during some well known periods of elevated volatility including the Great Depression, the technology boom of the late 1990s, and the Great Recession of the late 2000s (CGP, 2013).

The Online Appendix sheds light on the drivers of the difference between the positive volatility beta that we estimate for the market as a whole in the modern period and the negative volatility beta that BKSY estimate. While we confirm the result that in the BKSY (2014) specification market innovations are negatively correlated with N_V, that result is sensitive to the exact specification. If *RVAR* is computed using daily instead of monthly returns, in particular, the correlation moves much closer to zero and in several cases becomes positive, as in our baseline specification.

One important driver of the correlation between market returns and N_V is the correlation between N_{DR} and N_V. Because an increase in discount rates lowers stock prices, other things equal, these two correlations tend to have opposite signs. In our replication of BKSY's analysis, we find a positive correlation of 0.47 between N_{DR} and N_V, but this positive correlation does not survive if quarterly data are used instead of yearly data, if the value spread is used in the VAR, or if *RVAR* is constructed using daily instead of monthly returns. In all these alternative cases, the relation between N_{DR} and N_V is much weaker or even negative, confirming the results of a long literature in asset pricing (see, e.g., Lettau and Ludvigson, 2010).

In summary, we believe that neither the finding of a negative volatility beta for value stocks relative to growth stocks nor the finding of a positive volatility beta for the aggregate equity market in the modern period should be surprising. Stockholders are long options, both options to invest in growth opportunities (particularly important for growth firms) and options to default on bondholders. These options become more valuable when volatility increases, driving up stock prices. Thus, no theoretical reason exists to believe that higher volatility always reduces aggregate stock prices. And, in recent history, there have been important episodes in which stock prices have been both high and volatile, most notably in the stock boom of the 1990s.

[21] BKSY estimate their VAR system by GMM, using additional moment conditions implied by the ICAPM and the unconditional returns on test assets. We used a similar methodology for a two-beta ICAPM model in Campbell et al. (2013) but found it to be computationally challenging and numerically unstable. We have not replicated this approach for the three-beta ICAPM, but we do not believe it has a first-order effect on the differences in empirical results because we can account for these differences using unrestricted VAR models.

7.2. Comparison with consumption-based models

In this paper, as in Campbell (1993), we have estimated our model without having to observe the consumption process of the investor (who was assumed to hold the market portfolio). However, the model could also be expressed in terms of the investor's consumption. Both consumption and asset returns are endogenous, and the two representations are equivalent.

In this subsection we show how to map the returns-based representation to the consumption-based representation. We focus on two main objects of interest: consumption innovations and the stochastic discount factor.

Consumption innovations for our investor are given by

$$\Delta c_{t+1} - E_t \Delta c_{t+1} = (r_{t+1} - E_t r_{t+1}) - (\psi - 1) N_{DR,t+1} - (\psi - 1)\frac{1}{2}\frac{\omega}{1-\gamma} N_{V,t+1}. \quad (25)$$

The EIS parameter ψ, which enters this equation, is not pinned down by our VAR estimation or the cross section of risk premia, so we calibrate it to three different values: 0.5, 1.0, and 1.5. The Online Appendix shows that implied consumption volatility is positively related to ψ, given our VAR estimates of return dynamics. With $\psi = 0.5$, our investor's consumption (which need not equal aggregate consumption) is considerably more volatile than aggregate consumption but roughly as volatile as the time series of stockholder's consumption we obtained from Malloy et al. (2009). Implied and actual consumption growth are positively correlated, and a stockholder's consumption correlates with implied consumption more strongly than aggregate consumption.

We can also represent the entire SDF in terms of consumption. We can write it as a function of consumption innovations ($\Delta c_{t+1} - E_t \Delta c_{t+1}$), news about future consumption growth (N_{CF}), and news about future consumption volatility ($N_{CV,t+1}$):

$$m_{t+1} - E_t m_{t+1} = -\frac{1}{\psi}(\Delta c_{t+1} - E_t \Delta c_{t+1}) - (\gamma - \frac{1}{\psi}) N_{CF,t+1} + \frac{1}{2}\eta\left(\frac{\theta-1}{\theta}\right) N_{CV,t+1}, \quad (26)$$

where the parameter η is a constant that depends on the VAR parameters and on the structural parameters of the model (the Online Appendix reports the derivation). As in the case of the consumption innovations, the SDF depends on the parameter ψ. That parameter is not pinned down by risk premia in this model, thus requiring additional moments to be identified relative to our returns-based analysis.

This SDF corresponds to the standard SDF used in the consumption-based long-run risk literature (e.g., Bansal and Yaron, 2004). When $\gamma > \frac{1}{\psi}$, news about low future consumption growth or high volatility increases the investor's marginal utility, so assets that have low returns when such bad news arrives command an additional risk premium. The SDF collapses to the standard consumption-CAPM with power utility when $\gamma = \frac{1}{\psi}$ (and, therefore, $\theta = 1$). In that case, the coefficient on consumption innovation is simply equal to γ, and both the consumption news term and the volatility news term disappear from the SDF.

To conclude, the model can be equivalently expressed in terms of consumption or returns. In this paper, we follow Campbell (1993) using the latter approach, but we emphasize that neither approach is more structural than the other, as all quantities are determined jointly in equilibrium.

7.3. Implications for the risk-free rate

In addition to deriving the implied consumption process, we can use the estimated VAR and preference parameters to back out the implied risk-free rate in the economy, showing what time series for the risk-free rate would have made the long-run investor content not to time the market at each point in time.

In the Online Appendix, we show that the implied risk-free rate is the difference between the expected return on the market (which can be directly obtained from the VAR) and the market risk premium, itself a function of σ_t^2:

$$r_{t+1}^f = E_t r_{t+1}^M - H\sigma_t^2, \tag{27}$$

for a constant H that, in our data, is estimated to be 2.27. The implied risk-free rate therefore decreases (and potentially becomes negative) whenever conditional variance increases without a corresponding increase in the conditional expectation of the market return.

The Online Appendix shows that the implied risk-free rate is volatile (with a standard deviation of 2.4% per quarter). It became negative during the Great Depression, the technology boom, and the global financial crisis, all periods of elevated volatility. The implied risk-free rate therefore does not resemble the observed Treasury bill rate. This result should be expected. As discussed in Section 3.2.3, we do not impose the conditional implications of the model for the market risk premium, precisely because market volatility and expected market returns do not line up well in the data. For this reason, our model does not explain why a conservative long-term investor would not use Treasury bills as part of an equity market timing strategy.

The Online Appendix also shows that news about the present value of future implied risk-free rates has a volatility similar to that of news about market discount rates. Implied risk-free rate news was persistently negative during the Great Depression and the technology boom but not during the global financial crisis, which had a more transitory effect on the state variables of our model.

7.4. Robustness to empirical methodology

The Online Appendix examines the robustness of our results to a wide variety of methodological changes. We use various subsets of variables in our baseline VAR, estimate the VAR in different ways, use different estimates of realized variance, alter the set of variables in the VAR, explore the VAR's out-of-sample and split-sample properties, and use different proxies for the wealth portfolio including delevered equity portfolios. Such robustness analysis is important because the VAR's news decomposition can be sensitive to the forecasting variables included.[22]

Key results from these robustness tests follow. We find that including two of *DEF*, *PE*, and *VS* is generally essential for our finding of a negative β_V for *HML*. However, successful pricing by our volatility ICAPM requires all three in the VAR. We find a negative β_V for *HML* regardless of how we estimate the VAR (e.g., OLS or various forms of WLS) or construct our proxy for *RVAR*. However, our ICAPM is most successful at pricing using a quarterly VAR estimated using WLS, where *RVAR* is constructed from daily returns.

We also augment the set of variables under consideration to be included in the VAR. We not only explore different ways to measure the market's valuation ratio but also include other variables known to forecast aggregate returns and market volatility, specifically Lettau and Ludvigson's (2001) *CAY* variable and our quarterly *FIGARCH* forecast. *HML*'s β_V is always negative, and our volatility ICAPM generally does well in describing cross-sectional variation in average returns. We further find that our results are robust to using alternative proxies for the market portfolio, formed by combining Treasury Bills and the market in various constant proportions.

An important concern is the extent to which our VAR coefficients are stable over time. We address this issue in two ways. First, we generate the model's news terms out-of-sample, by estimating the VAR over an expanding window. We start the out-of-sample analysis in July 1963. We continue to find a negative β_V for *HML*, relative to our baseline result, and the cross-sectional R^2 increases to 77%. Second, we instead allow for a structural break between the early and modern periods in the coefficients of the return and volatility regressions of the VAR. We again find that *HML*'s β_V is negative. As with our baseline specification, the modern period cross-sectional R^2 is approximately 48%.

Finally, the Online Appendix describes in detail the results of analysis studying the volatility betas we have estimated for the market as a whole and for value stocks relative to growth stocks. For example, we report OLS estimates of simple betas on *RVAR* and the 15-year horizon *FIGARCH* forecast (FIG_{60}) for *HML* and *RMRF*. The betas based on these two simple proxies have the same sign as those using volatility news from our VAR.

8. Conclusion

We extend the approximate closed-form intertemporal capital asset pricing model of Campbell (1993) to allow for stochastic volatility. Our model recognizes that an investor's investment opportunities can deteriorate either because expected stock returns decline or because the volatility of stock returns increases. A long-term investor with Epstein-Zin preferences and relative risk aversion greater than one, holding an aggregate stock index,

[22] All our VAR systems forecast returns, not cash flows. As Engsted et al. (2012) clarify, results are approximately invariant to this decision, notwithstanding the concerns of Chen and Zhao (2009).

wishes to hedge against both types of changes in investment opportunities. Such an investor's perception of a stock's risk is determined not only by its beta with unexpected market returns and news about future returns (or, equivalently, news about market cash flows and discount rates) but also by its beta with news about future market volatility. Although our model has three dimensions of risk, the prices of all these risks are determined by a single free parameter, the investor's coefficient of relative risk aversion.

Our implementation models the return on the aggregate stock market as one element of a vector autoregressive system. The volatility of all shocks to the VAR is another element of the system. The estimated VAR system reveals new low-frequency movements in market volatility tied to the default spread. We show that the negative post-1963 CAPM alphas of growth stocks are justified because these stocks hedge long-term investors against both declining expected stock returns and increasing volatility. The addition of volatility risk to the model helps it fit the cross section of value and growth stocks, and small and large stocks, with a moderate, economically reasonable value of risk aversion.

We confront our model with portfolios of stocks sorted by past betas with the market return and volatility and portfolios double-sorted by characteristics and past volatility betas. We also confront our model with managed portfolios that vary equity exposure in response to our estimates of market variance. The explanatory power of the model is quite good across all these sets of test assets, with stable parameter estimates. Notably, the model helps to explain the low cross-sectional reward to past market beta and the negative return to idiosyncratic volatility as the result of volatility exposures of stocks with these characteristics in the post-1963 period.

Our model does not explain why a conservative long-term investor with constant risk aversion retains a constant equity exposure in response to changes in the equity premium that are not proportional to changes in the variance of stock returns. As a consequence, we do not interpret our model as a representative-agent model of general equilibrium in financial markets. However, our model does answer the interesting microeconomic question: Do reasonable preference parameters exist that would make a long-term investor, constrained to invest 100% in equity, content to hold the market instead of tilting toward value stocks or other high-return stock portfolios? Our answer is clearly yes.

References

Adrian, T., Rosenberg, J., 2008. Stock returns and volatility: pricing the short-run and long-run components of market risk. J. Financ. 63, 2997–3030.

Ait-Sahalia, Y., Karaman, M., Mancini, L., 2015. The Term Structure of Variance Swaps and Risk Premia. Princeton University, Princeton, NJ. Unpublished working paper.

Andersen, T., Bollerslev, T., Diebold, F., Labys, P., 2003. Modeling and forecasting realized volatility. Econometrica 71, 579–625.

Baillie, R., Bollerslev, T., Mikkelsen, H., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. J. Econom. 74, 3–30.

Bansal, R., Kiku, D., Shaliastovich, I., Yaron, A., 2014. Volatility, the macroeconomy and asset prices. J. Financ. 69, 2471–2511.

Bansal, R., Kiku, D., Yaron, A., 2012. An empirical evaluation of the long-run risks model for asset prices. Crit. Financ. Rev. 1, 183–221.

Bansal, R., Yaron, A., 2004. Risks for the long run. J. Financ. 59, 1481–1509.

Barndorff-Nielsen, O., Shephard, N., 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. J. R. Stat. Soc. 64, 253–280.

Beeler, J., Campbell, J., 2012. The long-run risks model and aggregate asset prices: an empirical assessment. Crit. Financ. Rev. 1, 141–182.

Berk, J., Green, R., Naik, V., 1999. Optimal investment, growth options, and security returns. J. Financ. 54, 1553–1607.

Black, F., 1976. Studies of stock price volatility changes. In: Proceedings of the 1976 Meetings of the American Statistical Association. Business and Economic Statistics Section, Washington, DC, pp. 177–181.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econom. 31, 307–327.

Breeden, D., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. J. Financ. Econ. 7, 265–296.

Calvet, L., Fisher, A., 2007. Multifrequency news and stock returns. J. Financ. Econ. 86, 178–212.

Campbell, J., 1987. Stock returns and the term structure. J. Financ. Econ. 18, 373–399.

Campbell, J., 1993. Intertemporal asset pricing without consumption data. Am. Econ. Rev. 83, 487–512.

Campbell, J., 1996. Understanding risk and return. J. Polit. Econ. 104, 298–345.

Campbell, J., Giglio, S., Polk, C., 2013. Hard times. Rev. Asset Pricing Stud. 3, 95–132.

Campbell, J., Hentschel, L., 1992. No news is good news: an asymmetric model of changing volatility in stock returns. J. Financ. Econ. 31, 281–318.

Campbell, J., Polk, C., Vuolteenaho, T., 2010. Growth or glamour? fundamentals and systematic risk in stock returns. Rev. Financ. Stud. 23, 305–344.

Campbell, J., Shiller, R., 1988a. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195–228.

Campbell, J., Shiller, R., 1988b. Stock prices, earnings, and expected dividends. J. Financ. 43, 661–676.

Campbell, J., Vuolteenaho, T., 2004. Bad beta, good beta. Am. Econ. Rev. 94, 1249–1275.

Carhart, M., 1997. On persistence in mutual fund performance. J. Financ. 52, 57–82.

Chen, J., 2003. Intertemporal CAPM and the cross section of stock returns. University of California, Davis, CA. Unpublished working paper.

Chen, L., Zhao, X., 2009. Return decomposition. Rev. Financ. Stud. 22, 5213–5249.

Christiansen, C., Schmeling, M., Schrimpf, A., 2012. A comprehensive look at financial volatility prediction by economic variables. J. Appl. Econom. 27, 956–977.

Christie, A., 1982. The stochastic behavior of common stock variances: value, leverage, and interest rate effects. J. Financ. Econ. 10, 407–432.

Coval, J., Shumway, T., 2001. Expected option returns. J. Financ. 66, 983–1009.

Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross-sectional variation in common stock returns. J. Financ. 52, 1–33.

Daniel, K., Titman, S., 2012. Testing factor-model explanations of market anomalies. Crit. Financ. Rev. 1, 103–139.

Davis, J., Fama, E., French, K., 2000. Characteristics, covariances, and average returns: 1929 to 1997. J. Financ. 55, 389–406.

Dew-Becker, I., Giglio, S., Le, A., Rodriguez, M., 2016. The price of variance risk. University of Chicago, Chicago, IL. Unpublished working paper.

Dou, W., 2016. Embrace or fear uncertainty: growth options, limited risk sharing, and asset prices. Massachusetts Institute of Technology, Cambridge, MA. Unpublished working paper.

Engle, R., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica 50, 987–1007.

Engle, R., Ghysels, E., Sohn, B., 2013. Stock market volatility and macroeconomic fundamentals. Rev. Econ. Stat. 95, 776–797.

Engsted, T., Pedersen, T., Tanggaard, C., 2012. Pitfalls in VAR based return decompositions: a clarification. J. Bank. Financ. 36, 1255–1265.

Epstein, L., Farhi, E., Strzalecki, T., 2014. How much would you pay to resolve long-run risk? Am. Econ. Rev. 104, 2680–2697.

Epstein, L., Zin, S., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica 57, 937–969.

Epstein, L., Zin, S., 1991. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: an empirical analysis. J. Polit. Econ. 99, 263–286.

Eraker, B., 2008. Affine general equilibrium models. Manag. Sci. 54, 2068–2080.
Fama, E., French, K., 1989. Business conditions and expected returns on stocks and bonds. J. Financ. Econ. 25, 23–50.
Fama, E., French, K., 1992. The cross section of expected stock returns. J. Financ. 47, 427–465.
Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56.
Fama, E., French, K., 2016. Dissecting anomalies with a five-factor model. Rev. Financ. Stud. 29, 69–103.
Frazzini, A., Pedersen, L., 2013. Betting against beta. J. Financ. Econ. 111, 1–25.
French, K., Schwert, W., Stambaugh, R., 1987. Expected stock returns and volatility. J. Financ. Econ. 19, 3–29.
Garcia, R., Renault, E., Semenov, A., 2006. Disentangling risk aversion and intertemporal substitution. Financ. Res. Lett. 3, 181–193.
Hansen, L., 2012. Dynamic valuation decomposition within stochastic economies. Econometrica 80, 911–967.
Hansen, L., Heaton, J., Lee, J., Roussanov, N., Heckman, J., Leamer, E., 2007. Intertemporal substitution and risk aversion. In: Handbook of Econometrics, Vol. 6A. North-Holland, Amsterdam, pp. 3967–4056.
Hansen, L., Heaton, J., Li, N., 2008. Consumption strikes back? measuring long-run risk. J. Polit. Econ. 116, 260–302.
Harvey, C., 1989. Time-varying conditional covariances in tests of asset pricing models. J. Financ. Econ. 24, 289–317.
Harvey, C., 1991. The world price of covariance risk. J. Financ. 46, 111–157.
Heston, S., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. 6, 327–343.
Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross section of expected returns. J. Financ. 51, 3–54.
Kandel, S., Stambaugh, R., 1991. Asset returns and intertemporal preferences. J. Monet. Econ. 27, 39–71.
Krishnamurthy, A., Vissing-Jørgensen, A., 2012. The aggregate demand for treasury debt. J. Polit. Econ. 120, 233–267.
Lettau, M., Ludvigson, S., 2001. Consumption, aggregate wealth, and expected stock returns. J. Financ. 56, 815–849.
Lettau, M., Ludvigson, S., 2010. Measuring and modeling variation in the risk-return trade-off. In: Ait-Sahalia, Y., Hansen, L. (Eds.), Handbook of Financial Econometrics. Elsevier, Amsterdam, pp. 617–690.
Lewellen, J., Nagel, S., Shanken, J., 2010. A skeptical appraisal of asset pricing tests. J. Financ. Econ. 96, 175–194.
Lustig, H., Roussanov, N., Verdelhan, A., 2011. Common risk factors in currency markets. Rev. Financ. Stud. 24, 3731–3777.
Lustig, H., Van Nieuwerburgh, S., Verdelhan, A., 2013. The wealth-consumption ratio. Rev. Asset Pricing Stud. 3, 38–94.
Malloy, C., Moskowitz, T., Vissing-Jø rgensen, A., 2009. Long-run stockholder consumption risk and asset returns. J. Financ. 64, 2427–2479.
McQuade, T., 2012. Stochastic volatility and asset pricing puzzles. Harvard University. Unpublished working paper.
Merton, R., 1973. An intertemporal capital asset pricing model. Econometrica 41, 867–887.
Nagel, S., 2016. The liquidity premium of near-money assets. Q. J. Econ. 131, 1927–1971.
Paye, B., 2012. Deja vol: predictive regressions for aggregate stock market volatility using macroeconomic variables. J. Financ. Econ. 106, 527–546.
Restoy, F., Weil, P., 1998. Approximate equilibrium asset prices. National Bureau of Economic Research, Cambridge, MA. Unpublished working paper.
Restoy, F., Weil, P., 2011. Approximate equilibrium asset prices. Rev. Fin. 15, 1–28.
Schwert, W., 1989. Why does stock market volatility change over time? J. Financ. 44, 1115–1153.
Sohn, B., 2010. Stock market volatility and trading strategy-based factors. Georgetown University, Washington, DC. Unpublished working paper.
Tallarini, T., 2000. Risk sensitive real business cycles. J. Monetary Econ. 45, 507–532.