An Intertemporal CAPM with Stochastic Volatility

John Y. Campbell, Stefano Giglio, Christopher Polk, and Robert Turley[1]

First draft: October 2011
This version: April 2014

[1]Campbell: Department of Economics, Littauer Center, Harvard University, Cambridge MA 02138, and NBER. Email john_campbell@harvard.edu. Phone 617-496-6448. Giglio: Booth School of Business, University of Chicago, 5807 S. Woodlawn Ave, Chicago IL 60637. Email stefano.giglio@chicagobooth.edu. Polk: Department of Finance, London School of Economics, London WC2A 2AE, UK. Email c.polk@lse.ac.uk. Turley: Baker Library 220D, Harvard Business School, Boston MA 02163. Email turley@fas.harvard.edu.
We are grateful to Torben Andersen, Gurdip Bakshi, John Cochrane, Bjorn Eraker, Bryan Kelly, Ian Martin, Sydney Ludvigson, Monika Piazzesi, Tuomo Vuolteenaho, and seminar participants at BI Norwegian School of Management, CEMFI, the 2012 EFA, 2011 HBS Finance Unit research retreat, IESE, INSEAD, 2012 Spring NBER Asset Pricing Meeting, Norges Bank Investment Management, Northwestern University, NYU, Oxford University, Stanford University, Swiss Finance Institute-Geneva, University of California-Berkeley, University of Chicago, University of Edinburgh, University of Illinois, University of Paris Dauphine, University of Pennsylvania, and the 2012 WFA for comments. We thank Josh Coval, Ken French, Nick Roussanov, Mila Getmansky Sherman, and Tyler Shumway for providing data used in the analysis.

Abstract

This paper extends the approximate closed-form intertemporal capital asset pricing model of Campbell (1993) to allow for stochastic volatility. The return on the aggregate stock market is modeled as one element of a vector autoregressive (VAR) system, and the volatility of all shocks to the VAR is another element of the system. Our estimates of this VAR reveal novel low-frequency movements in market volatility tied to the default spread. We show that growth stocks underperform value stocks because they hedge two types of deterioration in investment opportunities: declining expected stock returns, and increasing volatility. Our model also prices portfolios of stocks sorted by risk exposures, and non-equity assets including equity options, corporate bonds, and carry-trade currency portfolios.

JEL classification: G12, N22

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, conservative long-term investors will seek to hold "intertemporal hedges", assets that perform well when investment opportunities deteriorate. Such assets should deliver lower average returns in equilibrium if they are priced from conservative long-term investors' first-order conditions.

Since the seminal work of Merton (1973) on the intertemporal capital asset pricing model (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 (1988a) and the assumption that a representative investor has Epstein-Zin utility (Epstein and Zin 1989) 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 (VAR) estimates of asset return dynamics (Campbell 1996). Campbell and Vuolteenaho (2004), Campbell, Polk, and Vuolteenaho (2010), and Campbell, Giglio, and Polk (2013) use this approach to argue that value stocks outperform growth stocks on average because growth stocks do well when the expected return on the aggregate stock market declines; in other words, growth stocks have low risk premia because they are intertemporal hedges for long-term investors.

A weakness of the papers cited above is that they ignore time-variation in the volatility of stock returns. In general, investment opportunities may deteriorate either because expected stock returns decline or because the volatility of stock returns increases, and it is an empirical question which of these two types of intertemporal risk have a greater effect on asset returns. We address this weakness in this paper by extending the approximate closed-form ICAPM to allow for stochastic volatility. The resulting model explains risk premia in the stock market using 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 the 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.

Since the long-horizon investor in our model cares mostly about persistent changes in the investment opportunity set, there must be predictable long-run variation in volatility for volatility risk to matter. Empirically, we implement our methodology using a vector autoregression (VAR) including stock returns, realized variance, and other financial indicators that may be relevant for predicting returns and risk. Our VAR reveals low-frequency movements in market volatility tied to the default spread, the yield spread of low-rated over high-rated bonds. While this phenomenon has received little attention in the literature, we argue that it is sensible: Investors in risky bonds perceive the long-run component of volatility and incorporate this information when they set credit spreads, as risky bonds are short the option to default over long maturities. Moreover, we show that GARCH-based methods that filter only the information in past returns in order to disentangle the short-run and long-run volatility components miss this important low-frequency component.

With our novel model of long-run volatility in hand, we find that growth stocks have low average returns because they outperform not only when the expected stock return declines, but also when stock market volatility increases. Thus growth stocks hedge two types of deterioration in investment opportunities, not just one. In the period since 1963 that creates the greatest empirical difficulties for the standard CAPM, we find that the three-beta model explains over 62% of the cross-sectional variation in average returns of 25 portfolios sorted by size and book-to-market ratios. The model is not rejected at the 5% level while the CAPM is strongly rejected. The implied coefficient of relative risk aversion is an economically reasonable 6.9, in contrast to the much larger estimate of 20.7, which we get when we estimate a comparable version of the two-beta ICAPM of Campbell and Vuolteenaho (2004) using the same data.[2] This success is due in large part to the inclusion of volatility betas in the specification. The cross-sectional spread in volatility betas generates an annualized average return spread of 5.2% compared to spreads of 2.8% and 2.2% for cash-flow and discount-rate betas.

We confirm that our findings are robust by looking at five alternative sets of test assets. First we sort stocks by their estimated risks, specifically their past market and volatility betas, and create risk-sorted portfolios. We do this to respond to the argument of Daniel and Titman (1997, 2012) and Lewellen, Nagel, and Shanken (2010) that asset-pricing tests using only portfolios sorted by characteristics known to be related to average returns, such as size and value, can be misleading when the resulting portfolios have a low-dimensional factor structure. Second, we sort stocks by two characteristics—their book-market ratios and their idiosyncratic volatility—as well as their past volatility betas. This uncovers interesting interactions among these characteristics and volatility betas, and allows us to address the idiosyncratic volatility effect of Ang, Hodrick, Xing, and Zhang (2006), which is a major challenge to extant asset-pricing models. Third, we look at option and bond portfolios that are exposed to aggregate volatility risk, specifically the S&P 100 index straddle of Coval and Shumway (2001) and the risky bond factor of Fama and French (1993), which should be sensitive to changes in aggregate volatility since risky corporate debt is short the option to default. Fourth, we look at currency portfolios sorted by interest rates in the manner of Lustig, Roussanov, and Verdelhan (2011), which capture the risk and return properties of the foreign exchange carry trade. Finally, we study variance forwards. These positions are exposed to realized volatility occurring from one to 12 months forward and thus present an important check on the internal consistency of our model. In all these cases we find that our three-beta ICAPM, unlike the CAPM or the two-beta ICAPM, delivers satisfactory pricing

[2] The risk aversion estimate reported in Campbell and Vuolteenaho's (2004) paper is 28.8.

results with minimal variation in parameters across test assets.

The organization of our paper is as follows. Section 2 reviews related literature. Section 3 lays out the approximate closed-form ICAPM and extends it to incorporate stochastic volatility. Section 4 presents data, econometrics, and VAR estimates of the dynamic process for stock returns and realized volatility. This section documents the empirical success of our model in forecasting long-run volatility. Section 5 introduces our test assets and estimates their betas with news about the market's future cash flows, discount rates, and volatility. Section 6 turns to cross-sectional asset pricing and estimates a representative investor's preference parameters to fit a cross-section of test assets, taking the dynamics of stock returns as given. Section 7 explores the implications of our model for the history of investors' marginal utility and consumption behavior. Section 8 presents a set of robustness exercises in which we vary our basic VAR specification for the dynamics of aggregate returns and risk, and explore the underlying components of volatility betas for the market portfolio and for value stocks versus growth stocks. Section 9 concludes. An online appendix to the paper (Campbell, Giglio, Polk, and Turley 2014) provides supporting details.

2 Literature Review

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 representative 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) pointed 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) used the same logic to substitute out the market portfolio return, turning the model into a generalized consumption CAPM in the style of Breeden (1979).

Kandel and Stambaugh (1991) were the first researchers to study the implications for asset returns of time-varying first and second moments of consumption growth in a model with a representative Epstein-Zin investor. Specifically, Kandel and Stambaugh (1991) assumed a four-state Markov chain for the expected growth rate and conditional volatility of consumption, and provided closed-form solutions for important asset-pricing moments. In the same spirit Bansal and Yaron (2004) added stochastic volatility to the Restoy-Weil model, and subsequent theoretical and empirical research in the long-run risk framework has increasingly emphasized the importance of stochastic volatility (Bansal, Kiku, and Yaron 2012, Beeler and Campbell 2012, Hansen 2012). In this paper we give the approximate closed-form ICAPM the same capability to handle stochastic volatility that its cousin, the long-run risk model, already possesses.

One might ask whether there is any reason to work with an ICAPM rather than a consumption-based model given that these are two representations of the same underlying economic structure. Writing the model as an ICAPM is conceptually appealing because it describes risks as they appear to an investor who takes asset prices as given and chooses consumption to satisfy his budget constraint. This is a valuable microeconomic complement to the macroeconomic perspective provided by the consumption-based approach.

The ICAPM also has several advantages as an empirical specification. First, it does not require that all agents in the economy participate in financial markets. So long as an agent with Epstein-Zin preferences holds the aggregate wealth portfolio, the ICAPM will hold even if there are other agents in the economy who receive nonfinancial income and account for some portion of aggregate consumption.[3] Related to this, the ICAPM can be given a microeconomic interpretation as a description of the conditions that are required for a long-term equity investor not to tilt his portfolio towards stocks with high average returns such as value stocks. We discuss this interpretation briefly in the conclusion of the paper.

Second, the ICAPM allows an empirical analysis based on financial proxies for the aggregate market portfolio rather than on accurate measurement of aggregate consumption. While there are certainly challenges to the accurate measurement of financial wealth, financial time series are generally available on a more timely basis and over longer sample periods than consumption series. Our analysis follows a large literature that assumes that the return on a broad stock index is a reasonable proxy for the return on aggregate wealth. In the robustness section, we explore the sensitivity of our results to using other proxies for the wealth portfolio.

Third, the ICAPM generates empirical predictions that depend on the coefficient of relative risk aversion but not the elasticity of intertemporal substitution. This means that we do not need to assume a wedge between risk aversion and the reciprocal of the elasticity of intertemporal substitution, and therefore do not face the critique of Epstein, Farhi, and Strzalecki (2013) that a large wedge implies an unrealistic willingness to pay for early resolution of uncertainty.[4]

The particular specification of the ICAPM in this paper has two further advantages. It is flexible enough to allow multiple state variables that can be estimated in a VAR system, so it does not require low-dimensional calibration of the sort used in the long-run risk literature. And we use an affine stochastic volatility process that governs the volatility of all state variables, including itself. We show that this assumption fits financial data reasonably well, and it guarantees that stochastic volatility would always remain positive in a continuous-

[3] Hansen (2014) highlights this point in his Nobel Lecture. Malloy, Moskowitz, and Vissing-Jørgensen (2009) study long-run consumption risk using microeconomic data on stockholders' consumption growth, but they can only do this over a relatively short sample period.

[4] 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, although Garcia, Renault, and Semenov (2006) and Hansen, Heaton, Lee, and Roussanov (2007) point out that this interpretation may not be correct when γ differs from the reciprocal of ψ.

time version of the model, a property that does not always hold in current implementations of the long-run risk model.[5]

The closest precursors to our work are unpublished papers by Chen (2003) and Sohn (2010). Both papers explore the effects of stochastic volatility on asset prices in an ICAPM setting but make strong assumptions about the covariance structure of various news terms when deriving their pricing equations. Chen (2003) assumes constant covariances between shocks to the market return (and powers of those shocks) and news about future expected market return variance. Sohn (2010) makes two strong assumptions about asset returns and consumption growth, specifically that all assets have zero covariance with news about future consumption growth volatility and that the conditional contemporaneous correlation between the market return and consumption growth is constant through time. Duffee (2005) presents evidence against the latter assumption. It is in any case unattractive to make assumptions about consumption growth in an ICAPM that does not require accurate measurement of consumption.

Chen estimates a VAR with a GARCH model to allow for time variation in the volatility of return shocks, restricting market volatility to depend only on its past realizations and not those of the other state variables. His empirical analysis has little success in explaining the cross-section of stock returns. Sohn uses a similar but more sophisticated GARCH model for market volatility and tests how well short-run and long-run risk components from the GARCH estimation can explain the returns of various stock portfolios, comparing the results to factors previously shown to be empirically successful. In contrast, our paper incorporates the volatility process directly in the ICAPM, allowing heteroskedasticity to affect and to be predicted by all state variables, and showing how the price of volatility risk is pinned down by the time-series structure of the model along with the investor's coefficient of risk aversion.

Bansal, Kiku, Shaliastovich and Yaron (2013), a paper contemporaneous with ours, explores the effects of stochastic volatility in the long-run risk model. Like us, they find stochastic volatility to be an important feature in the time series of equity returns. There are some important differences in the underlying models. In particular, in their benchmark model they assume that the stochastic process driving volatility is homoskedastic. In our theoretical analysis we discuss some conditions that are required for their model solution to be valid, and argue that these conditions are not satisfied empirically. The different modeling assumptions and some differences in empirical implementation account for our contrasting empirical results; we show that volatility risk is very important in explaining the cross-section of stock returns while they find it has little impact on cross-sectional differences in risk premia. Indeed, Bansal, Kiku, Shaliastovich and Yaron (2013) find that a value-minus-growth bet has a positive beta with volatility news, while we find it always has a negative volatility

[5]Eraker (2008), Eraker and Shaliastovich (2008), and Heaton (2012) are exceptions whose models do guarantee positive volatility. Affine stochastic volatility models date back at least to Heston (1993) in continuous time, and have been developed and discussed by Ghysels, Harvey, and Renault (1996), Meddahi and Renault (2004), and Darolles, Gourieroux, and Jasiak (2006) among others.

beta. Our negative volatility beta estimate is more consistent with models of real options held by growth firms, such as McQuade (2012), and with the underperformance of value stocks during periods of elevated volatility including the Great Depression, the technology boom of the late 1990s, and the Great Recession of the late 2000s (Campbell, Giglio, and Polk 2013).

Stochastic volatility has been explored in other branches of the finance literature. For example, Chacko and Viceira (2005) and Liu (2007) show how stochastic volatility affects the optimal portfolio choice of long-term investors. Chacko and Viceira assume an AR(1) process for volatility and argue that movements in volatility are not persistent enough to generate large intertemporal hedging demands. Our more flexible multivariate process does allow us to detect persistent long-run variation in volatility. Campbell and Hentschel (1992), Calvet and Fisher (2007), and Eraker and Wang (2011) argue that volatility shocks will lower aggregate stock prices by increasing expected returns, if they do not affect cash flows. The strength of this volatility feedback effect depends on the persistence of the volatility process. Coval and Shumway (2001), Ang, Hodrick, Xing, and Zhang (2006), and Adrian and Rosenberg (2008) present evidence that shocks to market volatility are priced risk factors in the cross-section of stock returns, but they do not develop any theory to explain the risk prices for these factors.

Time-varying volatility is a prime concern of the field of financial econometrics. Since Engle's (1982) seminal paper on ARCH, much of the financial econometrics literature has focused on variants of the univariate GARCH model (Bollerslev 1986), in which return volatility is modeled as a function of past shocks to returns and of its own lags (see Poon and Granger (2003) and Andersen et al. (2006) for recent surveys). More recently, realized volatility from high-frequency data has been used to estimate stochastic volatility processes (Barndorff-Nielsen and Shephard 2002, Andersen et al. 2003). The use of realized volatility has improved the modeling and forecasting of volatility, including its long-run component; however, this literature has primarily focused on the information content of high-frequency intra-daily return data. This allows very precise measurement of volatility, but at the same time, given data availability constraints, limits the potential to use long time series to learn about long-run movements in volatility. In our paper, we measure realized volatility only with daily data, but augment this information with other financial time series that reveal information investors have about underlying volatility components.

A much smaller literature has, like us, looked directly at the information in other variables concerning future volatility. In early work, Schwert (1989) links movements in stock market volatility to various indicators of economic activity, particularly the price-earnings ratio and the default spread, but finds relatively weak connections. Engle, Ghysels and Sohn (2013) study the effect of inflation and industrial production growth on volatility, finding a significant link between the two, especially at long horizons. Campbell and Taksler (2003) look at the cross-sectional link between corporate bond yields and equity volatility, emphasizing that bond yields respond to idiosyncratic firm-level volatility as well as aggregate volatility. Two recent papers, Paye (2012) and Christiansen et al. (2012), look at larger sets of potential volatility predictors, including the default spread and valuation ratios, to find those that have predictive power for quarterly realized variance. The former paper, in a standard regression framework, finds that the commercial paper to Treasury spread and the default spread, among other variables, contain useful information for predicting volatility. The latter uses Bayesian Model Averaging to find the most successful predictors, and documents the importance of the default spread and valuation ratios in forecasting short-run volatility.

Finally, our paper relates to the enormous literature on cross-sectional patterns in stock returns. Most obviously, we contribute to the longstanding debate on the explanation of the value effect. Beyond this, our alternative test assets allow us to make two other contributions. By sorting stocks on their estimated market and volatility betas, we show that in the period since 1963, stocks with high past market betas tend to have high future volatility betas, a desirable property which reduces their equilibrium returns and may help to explain the weak relation between past market betas and returns (Fama and French 1992). Also, by sorting stocks on book-market ratios, idiosyncratic volatility, and past volatility betas, we show that idiosyncratic volatility is associated with higher volatility betas among growth stocks but much less so among value stocks. This is consistent with the intuition that idiosyncratic volatility reveals the presence of growth options among growth stocks, and it may also help to explain the finding of Ang, Hodrick, Xing, and Zhang (2006) that idiosyncratic volatility is negatively associated with average returns.[6]

3 An Intertemporal Model with Stochastic Volatility

3.1 Asset pricing with time varying risk

Preferences

We begin by assuming a representative agent with Epstein-Zin preferences. We write the 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

[6] Ang, Hodrick, Xing, and Zhang (2006) argue that the idiosyncratic volatility effect cannot be explained by aggregate volatility risk while Barinov (2013) and Chen and Petkova (2014) argue that it can. None of these papers develop a fully-fleshed-out theory to motivate their specific volatility risk factor or explain the risk price of that factor. The idiosyncratic volatility effect is non-monotonic as documented in Table VI of Ang, Hodrick, Xing, and Zhang (2006), and it varies across samples and methodologies as documented by Bali and Cakici (2009) and Fu (2009), possibly reflecting the interaction with the book-market ratio documented here.

γ, and the elasticity of intertemporal substitution ψ. For convenience, we define $\theta = (1-\gamma)/(1-1/\psi)$.

The corresponding stochastic discount factor (SDF) can be written as

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

where W_t is the market value of the consumption stream owned by the agent, including current consumption C_t.[7] The log return on wealth is $r_{t+1} = \ln\left(W_{t+1}/\left(W_t - C_t\right)\right)$, the log value of wealth tomorrow divided by reinvested wealth today. The log SDF is therefore

$$m_{t+1} = \theta \ln \delta - \frac{\theta}{\psi}\Delta c_{t+1} + (\theta - 1)\, r_{t+1}. \tag{3}$$

A convenient identity

The gross return to wealth can be written

$$1 + R_{t+1} = \frac{W_{t+1}}{W_t - C_t} = \left(\frac{C_t}{W_t - C_t}\right)\left(\frac{C_{t+1}}{C_t}\right)\left(\frac{W_{t+1}}{C_{t+1}}\right), \tag{4}$$

expressing it as the product of the current consumption payout, the growth in consumption, and the future price of a unit of consumption.

We find it convenient to work in logs. We define the log value of reinvested wealth per unit of consumption as $z_t = \ln\left(\left(W_t - C_t\right)/C_t\right)$, and the future value of a consumption claim as $h_{t+1} = \ln\left(W_{t+1}/C_{t+1}\right)$, so that the log return is:

$$r_{t+1} = -z_t + \Delta c_{t+1} + h_{t+1}. \tag{5}$$

Heuristically, the return on wealth is negatively related to the current value of reinvested wealth and positively related to consumption growth and the future value of wealth. The last term in equation (5) will capture the effects of intertemporal hedging on asset prices, hence the choice of the notation h_{t+1} for this term.

The ICAPM

We assume that asset returns are jointly conditionally lognormal, but we allow changing conditional volatility so we are careful to write second moments with time subscripts to indicate that they can vary over time. Under this standard assumption, the expected return on any asset must satisfy

$$0 = \ln \mathrm{E}_t \exp\{m_{t+1} + r_{i,t+1}\} = \mathrm{E}_t\left[m_{t+1} + r_{i,t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[m_{t+1} + r_{i,t+1}\right], \tag{6}$$

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

and the risk premium on any asset is given by

$$\mathrm{E}_t r_{i,t+1} - r_{f,t} + \frac{1}{2}\mathrm{Var}_t r_{t+1} = -\mathrm{Cov}_t\left[m_{t+1}, r_{i,t+1}\right]. \tag{7}$$

The convenient identity (5) can be used to write the log SDF (3) without reference to consumption growth:

$$m_{t+1} = \theta \ln \delta - \frac{\theta}{\psi} z_t + \frac{\theta}{\psi} h_{t+1} - \gamma r_{t+1}. \tag{8}$$

Since the first two terms in (5) are known at time t, only the latter two terms appear in the conditional covariance in (7). We obtain an ICAPM pricing equation that relates the risk premium on any asset to the asset's covariance with the wealth return and with shocks to future consumption claim values:

$$\mathrm{E}_t r_{i,t+1} - r_{f,t} + \frac{1}{2}\mathrm{Var}_t r_{t+1} = \gamma \mathrm{Cov}_t\left[r_{i,t+1}, r_{t+1}\right] - \frac{\theta}{\psi}\mathrm{Cov}_t\left[r_{i,t+1}, h_{t+1}\right] \tag{9}$$

Return and risk shocks in the ICAPM

To better understand the intertemporal hedging component h_{t+1}, we proceed in two steps. First, we approximate the relationship of h_{t+1} and z_{t+1} by taking a loglinear approximation about $\bar{z}$:

$$h_{t+1} \approx \kappa + \rho z_{t+1} \tag{10}$$

where the loglinearization parameter $\rho = \exp(\bar{z})/(1+\exp(\bar{z})) \approx 1 - C/W$.

Second, we apply the general pricing equation (6) to the wealth portfolio itself (setting $r_{i,t+1} = r_{t+1}$), and use the convenient identity (5) to substitute out consumption growth from this expression. Rearranging, we can write the variable z_t as

$$z_t = \psi \ln \delta + (\psi - 1)\mathrm{E}_t r_{t+1} + \mathrm{E}_t h_{t+1} + \frac{\psi}{\theta}\frac{1}{2}\mathrm{Var}_t\left[m_{t+1} + r_{t+1}\right]. \tag{11}$$

Third, we combine these expressions to obtain the innovation in h_{t+1}:

$$\begin{aligned} h_{t+1} - \mathrm{E}_t h_{t+1} &= \rho(z_{t+1} - \mathrm{E}_t z_{t+1}) \\ &= (\mathrm{E}_{t+1} - \mathrm{E}_t)\rho\left((\psi - 1)r_{t+2} + h_{t+2} + \frac{\psi}{\theta}\frac{1}{2}\mathrm{Var}_{t+1}\left[m_{t+2} + r_{t+2}\right]\right). \end{aligned} \tag{12}$$

Solving forward to an infinite horizon,

$$\begin{aligned} h_{t+1} - \mathrm{E}_t h_{t+1} &= (\psi - 1)(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j r_{t+1+j} \\ &\quad + \frac{1}{2}\frac{\psi}{\theta}(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j \mathrm{Var}_{t+j}\left[m_{t+1+j} + r_{t+1+j}\right] \\ &= (\psi - 1)N_{DR,t+1} + \frac{1}{2}\frac{\psi}{\theta}N_{RISK,t+1}. \end{aligned} \tag{13}$$

The second equality follows Campbell and Vuolteenaho (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}.

Finally, we substitute back into the intertemporal model (9):

$$\begin{aligned} & \mathrm{E}_t r_{i,t+1} - r_{f,t} + \frac{1}{2}\mathrm{Var}_t r_{i,t+1} \\ = \; & \gamma \mathrm{Cov}_t \left[r_{i,t+1}, r_{t+1} \right] + (\gamma - 1) \mathrm{Cov}_t \left[r_{i,t+1}, N_{DR,t+1} \right] - \frac{1}{2}\mathrm{Cov}_t \left[r_{i,t+1}, N_{RISK,t+1} \right] \\ = \; & \gamma \mathrm{Cov}_t \left[r_{i,t+1}, N_{CF,t+1} \right] + \mathrm{Cov}_t \left[r_{i,t+1}, -N_{DR,t+1} \right] - \frac{1}{2}\mathrm{Cov}_t \left[r_{i,t+1}, N_{RISK,t+1} \right]. \qquad (14) \end{aligned}$$

The first equality expresses the risk premium as risk aversion γ times covariance with the current market return, plus $(\gamma - 1)$ times covariance with news about future market returns, minus one half covariance with risk. This is an extension of the ICAPM as written by Campbell (1993), with no reference to consumption or the elasticity of intertemporal substitution ψ.[8] When the investor's risk aversion is greater than 1, assets which hedge aggregate discount rates ($\mathrm{Cov}_t \left[r_{i,t+1}, N_{DR,t+1} \right] < 0$) or aggregate risk ($\mathrm{Cov}_t \left[r_{i,t+1}, N_{RISK,t+1} \right] > 0$) have lower expected returns, all else equal.

The second equality rewrites the model, following Campbell and Vuolteenaho (2004), by breaking the market return into cash-flow news and discount-rate news. Cash-flow news N_{CF} is defined by $N_{CF} = r_{t+1} - \mathrm{E}_t r_{t+1} + N_{DR}$. The price of risk for cash-flow news is γ times greater than the price of risk for discount-rate news, hence Campbell and Vuolteenaho call betas with cash-flow news "bad betas" and those with discount-rate news "good betas". The third term in (14) shows the risk premium associated with exposure to news about future risks and did not appear in Campbell and Vuolteenaho's model, which assumed homoskedasticity. Not surprisingly, the coefficient is negative, indicating that an asset providing positive returns when risk expectations increase will offer a lower return on average.

While the elasticity of intertemporal substitution ψ does not affect risk premia in our model, this parameter does influence the implied behavior of the investor's consumption. We explore this further in the online appendix to the paper.

3.2 From risk to volatility

The risk shocks defined in the previous subsection are shocks to the conditional volatility of returns plus the stochastic discount factor, and therefore are not directly observable. We

[8] Campbell (1993) briefly considers the heteroskedastic case, noting that when $\gamma = 1$, $\mathrm{Var}_t \left[m_{t+1} + r_{t+1} \right]$ 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).

now make additional assumptions on the data generating process for stock returns that allow us to estimate the news terms. These assumptions imply that the conditional volatility of returns plus the stochastic discount factor is proportional to the conditional volatility of returns themselves.

Suppose the economy is described by a first-order VAR

$$\mathbf{x}_{t+1} = \bar{\mathbf{x}} + \mathbf{\Gamma}\left(\mathbf{x}_t - \bar{\mathbf{x}}\right) + \sigma_t \mathbf{u}_{t+1}, \tag{15}$$

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 $\Sigma_{11} = 1$.

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 volatility itself, have innovations whose variances move in proportion to one another. This assumption makes the stochastic volatility process affine, as in Heston (1993) and related work discussed above in our literature review.

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

$$\begin{aligned} N_{DR,t+1} &= \left(E_{t+1} - E_t\right) \sum_{j=1}^{\infty} \rho^j r_{t+1+j} \\ &= \mathbf{e}_1' \sum_{j=1}^{\infty} \rho^j \mathbf{\Gamma}^j \sigma_t \mathbf{u}_{t+1} \\ &= \mathbf{e}_1' \rho \mathbf{\Gamma} \left(\mathbf{I} - \rho\mathbf{\Gamma}\right)^{-1} \sigma_t \mathbf{u}_{t+1}, \end{aligned} \tag{16}$$

while implied cash flow news is:

$$\begin{aligned} N_{CF,t+1} &= \left(r_{t+1} - E_t r_{t+1}\right) + N_{DR,t+1} \\ &= \left(\mathbf{e}_1' + \mathbf{e}_1' \rho \mathbf{\Gamma}(\mathbf{I} - \rho\mathbf{\Gamma})^{-1}\right) \sigma_t \mathbf{u}_{t+1}. \end{aligned} \tag{17}$$

Furthermore, our log-linear model will make the log SDF, m_{t+1}, a linear function of the state variables. Since all shocks to the SDF are then proportional to σ_t, $\mathrm{Var}_t\left[m_{t+1} + r_{t+1}\right] \propto \sigma^2_t$. As a result, the conditional variance, $\mathrm{Var}_t\left[\left(m_{t+1} + r_{t+1}\right)/\sigma_t\right] = \omega_t$, will be a constant that does not depend on the state variables. Without knowing the parameters of the utility function, we can write $\mathrm{Var}_t\left[m_{t+1} + r_{t+1}\right] = \omega\sigma^2_t$ so that the news about risk, N_{RISK}, is proportional to news about market return variance, N_V.

$$
\begin{aligned}
N_{RISK,t+1} &= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \mathrm{Var}_{t+j} \left[r_{t+1+j} + m_{t+1+j} \right] \\
&= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \left(\omega \sigma_{t+j}^2 \right) \\
&= \omega \rho \mathbf{e}_2' \sum_{j=0}^{\infty} \rho^j \mathbf{\Gamma}^j \sigma_t \mathbf{u}_{t+1} \\
&= \omega \rho \mathbf{e}_2' \left(\mathbf{I} - \rho \mathbf{\Gamma} \right)^{-1} \sigma_t \mathbf{u}_{t+1} = \omega N_{V,t+1}. \qquad (18)
\end{aligned}
$$

Substituting (18) into (14), we obtain an empirically-testable intertemporal CAPM with stochastic volatility:

$$
\begin{aligned}
& \mathrm{E}_t r_{i,t+1} - r_{f,t} + \frac{1}{2} \mathrm{Var}_t r_{i,t+1} \\
= \; & \gamma \mathrm{Cov}_t \left[r_{i,t+1}, N_{CF,t+1} \right] + \mathrm{Cov}_t \left[r_{i,t+1}, -N_{DR,t+1} \right] - \frac{1}{2} \omega \mathrm{Cov}_t \left[r_{i,t+1}, N_{V,t+1} \right], \qquad (19)
\end{aligned}
$$

where covariances with news about three key attributes of the market portfolio (cash flows, discount rates, and volatility) describe the cross section of average returns.

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:

$$
\omega \sigma_t^2 = (1-\gamma)^2 \mathrm{Var}_t \left[N_{CF_{t+1}} \right] + \omega (1-\gamma) \mathrm{Cov}_t \left[N_{CF_{t+1}}, N_{V_{t+1}}, \right] + \omega^2 \frac{1}{4} \mathrm{Var}_t \left[N_{V_{t+1}} \right]. \qquad (20)
$$

We can see two main channels through which γ affects ω. First, a higher risk aversion—given the underlying volatilities of all shocks—implies a more volatile stochastic discount factor m, and therefore a higher RISK. This effect is proportional to $(1-\gamma)^2$, so it increases rapidly with γ. Second, there is a feedback effect on 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.[9]

[9] Bansal, Kiku, Shaliastovich and Yaron (2013) derive a similar expression, equation (16) in their paper. They claim that the equivalent expression for ω in their model reduces to $(1-\gamma)^2$ in the case of homoskedastic volatility (their equation 17). We discuss the conditions required for their claim to be valid in the next subsection. In a robustness test, they also derive a corresponding equation for the case of time-varying volatility (their equation C.4), but proceed with a linearization procedure that, as we discuss below, allows the parameters of the model to lie in a region of the parameter space where the true model does not have a solution.

This equation can also be written directly in terms of the VAR parameters. We define $\mathbf{x}_{CF}$ and $\mathbf{x}_V$ as the error-to-news vectors that map VAR innovations to volatility-scaled news terms:

$$\frac{1}{\sigma_t}N_{CF,t+1} = \mathbf{x}_{CF}\mathbf{u}_{t+1} = \left(\mathbf{e}_1' + \mathbf{e}_1'\rho\boldsymbol{\Gamma}(\mathbf{I}-\rho\boldsymbol{\Gamma})^{-1}\right)\mathbf{u}_{t+1} \tag{21}$$

$$\frac{1}{\sigma_t}N_{V,t+1} = \mathbf{x}_V\mathbf{u}_{t+1} = \left(\mathbf{e}_2'\rho(\mathbf{I}-\rho\boldsymbol{\Gamma})^{-1}\right)\mathbf{u}_{t+1}. \tag{22}$$

Then ω solves

$$0 = \omega^2\frac{1}{4}\mathbf{x}_V\boldsymbol{\Sigma}\mathbf{x}_V' - \omega\left(1-(1-\gamma)\,\mathbf{x}_{CF}\boldsymbol{\Sigma}\mathbf{x}_V'\right) + (1-\gamma)^2\,\mathbf{x}_{CF}\boldsymbol{\Sigma}\mathbf{x}_{CF}' \tag{23}$$

This quadratic equation for ω has two solutions, but the online appendix shows that one of the solutions can be disregarded. This false solution is easily identified by its implication that ω becomes infinite as volatility shocks become small. The correct solution is

$$\omega = \frac{1-(1-\gamma)\mathbf{x}_{CF}\boldsymbol{\Sigma}\mathbf{x}_V' - \sqrt{(1-(1-\gamma)\mathbf{x}_{CF}\boldsymbol{\Sigma}\mathbf{x}_V')^2 - (1-\gamma)^2(\mathbf{x}_V\boldsymbol{\Sigma}\mathbf{x}_V')(\mathbf{x}_{CF}\boldsymbol{\Sigma}\mathbf{x}_{CF}')}}{\frac{1}{2}\mathbf{x}_V\boldsymbol{\Sigma}\mathbf{x}_V'} \tag{24}$$

3.3 Conditions for the existence of a solution

Equation (23) has a real solution only if

$$(1-(1-\gamma)\mathbf{x}_{CF}\boldsymbol{\Sigma}\mathbf{x}_V')^2 - (1-\gamma)^2(\mathbf{x}_V\boldsymbol{\Sigma}\mathbf{x}_V')(\mathbf{x}_{CF}\boldsymbol{\Sigma}\mathbf{x}_{CF}') \geq 0. \tag{25}$$

Otherwise, if risk aversion, volatility shocks and cash flow shocks are large enough, as measured by the product $(1-\gamma)^2(\mathbf{x}_V\boldsymbol{\Sigma}\mathbf{x}_V')(\mathbf{x}_{CF}\boldsymbol{\Sigma}\mathbf{x}_{CF}')$, equation (23) may deliver a complex rather than a real value for ω. Given our VAR estimates of the variance and covariance terms, we find a real solution for ω as γ ranges from zero to 6.9. Figure 1 plots ω as a function of γ based on these estimates.

The online appendix shows that the condition for the existence of a real solution can be written in a simpler form as

$$(\rho_n - 1)(1-\gamma)\sigma_{cf}\sigma_v \leq 1, \tag{26}$$

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

To further develop the intuition behind these equations, in the online appendix we study a simple example in which the link between the existence to a solution for equation (23) and the existence of a value function for the representative agent can be shown analytically.

We do this in the special case of $\psi = 1$, since 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, Heaton, and Li 2008). In the example we find 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, equation (25). This result indicates that the possible non-existence of a solution to the quadratic equation for ω is a deep feature of the model, not an artefact of our loglinear approximation—which is not needed in the special case where $\psi = 1$. We also show that 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, then the value function ceases to exist.

This constraint on the parameters of the model is ignored in Bansal, Kiku, Shaliastovich, and Yaron (BKSY 2013), when they consider the case of time-varying volatility of volatility as a robustness test in Sections II.E and III.D. There, rather than imposing that ω and γ satisfy equation (23), they proceed to linearize equation (23) so that a solution for omega exists for all values of gamma. Therefore, they allow combinations of (γ,ω) for which equation (23) doesn't admit a real solution – in which case, as we show in the appendix, the true model doesn't have a solution.[10]

The online appendix also considers the benchmark specification of BKSY in which the volatility process is homoskedastic. In this case the term $\text{Var}_t(m_{t+1} + r_{t+1})$ is not in general proportional to σ_t^2, but depends on both σ_t^2 and σ_t. Therefore, N_{RISK} (news about future values of $\text{Var}_t(m_{t+1} + r_{t+1})$) is not in general proportional to N_V, so that N_V is not in general the right news term to use in cross-sectional pricing. The appendix shows that proportionality of N_{RISK} and N_V in the homoskedastic case can only be obtained with additional special assumptions not stated by BKSY: that N_{CF} and N_V are uncorrelated, and that the N_V shock only depends on innovations of state variables which are themselves homoskedastic. As both of these assumptions are strongly rejected by the data, we do not further consider the model with homoskedastic σ_t^2.

In summary, this section has shown that in our model, the existence of a real solution for ω is tightly linked with the existence of the value function. As a consequence, in our empirical analysis we take seriously the constraint implied by the quadratic equation (23), and require that our parameter estimates satisfy this constraint. Given the high average returns to risky assets in historical data, this means in practice that our estimate of risk aversion is often equal to the estimated upper bound of 6.9.

[10]In a previous version of our paper, we used also used this linearization to solve equation (23). For the reasons explained above, we now instead require the non-linearized quadratic equation (23) to have a real solution.

4 Predicting Aggregate Stock Returns and Volatility

4.1 State variables

Our full VAR specification of the vector $\mathbf{x}_{t+1}$ includes six state variables, five of which are the same as in Campbell, Giglio and Polk (2011). To those five variables, we add an estimate of conditional volatility. 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 Securities Prices (CRSP) value-weighted stock index and the log return on the Consumer Price Index. This is a standard proxy for the aggregate wealth portfolio, but in the robustness section 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 ($EVAR$). This variable 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 the N_V term described above. To construct $EVAR_t$, we proceed as follows. We first construct 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 then generates a series of predicted values for $RVAR$ at each time $t+1$, that 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 $t+1$:

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

Note that though we describe our methodology in a two-step fashion where we first estimate $EVAR$ and then use $EVAR$ in a VAR, this is only for interpretability. Indeed, this approach to modeling $EVAR$ can be considered a simple renormalization of equivalent results we would find from a VAR that included $RVAR$ directly.[11]

The third variable is the price-earnings ratio (PE) from Shiller (2000), constructed as the price of the S&P 500 index divided by a ten-year trailing moving average of aggregate earnings of companies in the S&P 500 index. Following Graham and Dodd (1934), Campbell and Shiller (1988b, 1998) advocate averaging earnings over several years to avoid temporary spikes in the price-earnings ratio caused by cyclical declines in earnings. We avoid any interpolation of earnings as well as lag the moving average by one quarter in order to ensure that all components of the time-t price-earnings ratio are contemporaneously observable by time t. The ratio is log transformed.

[11]Since we weight observations based on $RVAR$ in the first stage and then reweight observations using $EVAR$ in the second stage, our two-stage approach in practice is not exactly the same as a one-stage approach. However, Panel B of Table 11 shows that results from a $RVAR$-weighted single-step estimation are qualitatively very similar to those produced by our two-stage approach.

Fourth, the term yield spread (TY) is obtained from Global Financial Data. We compute the TY series as the difference between the log yield on the 10-Year US Constant Maturity Bond (IGUSA10D) and the log yield on the 3-Month US Treasury Bill (ITUSA3D).

Fifth, the small-stock value spread (VS) is constructed from data on the six "elementary" equity portfolios also obtained from Professor French's website. These elementary portfolios, which are constructed at the end of each June, are the intersections of two portfolios formed on size (market equity, ME) and three portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoint for year t is the median NYSE market equity at the end of June of year t. BE/ME for June of year t is the book equity for the last fiscal year end in $t-1$ divided by ME for December of $t-1$. The BE/ME breakpoints are the 30th and 70th NYSE percentiles.

At the end of June of year t, we construct the small-stock value spread as the difference between the $\ln(BE/ME)$ of the small high-book-to-market portfolio and the $\ln(BE/ME)$ of the small low-book-to-market portfolio, where BE and ME are measured at the end of December of year $t-1$. For months from July to May, the small-stock value spread is constructed by adding the cumulative log return (from the previous June) on the small low-book-to-market portfolio to, and subtracting the cumulative log return on the small high-book-to-market portfolio from, the end-of-June small-stock value spread. The construction of this series follows Campbell and Vuolteenaho (2004) closely.

The sixth variable in our VAR is the default spread (DEF), defined as the difference between the log yield on Moody's BAA and AAA bonds. The series is obtained from the Federal Reserve Bank of St. Louis. Campbell, Giglio and Polk (2011) add the default spread to the Campbell and Vuolteenaho (2004) VAR specification in part because that variable is known to track time-series variation in expected real returns on the market portfolio (Fama and French, 1989), but mostly because shocks to the default spread should to some degree reflect news about aggregate default probabilities. Of course, news about aggregate default probabilities should in turn reflect news about the market's future cash flows and volatility.

4.2 Short-run volatility estimation

In order 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), where the weight of each observation pair ($RVAR_{t+1}$, $\mathbf{x}_t$) is initially based on the time-t value of $(RVAR)^{-1}$. However, to ensure that the ratio of weights across observations is not extreme, we shrink these initial weights towards equal weights. In particular, 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 of the degree of variation over time in *expected* market return variance. More importantly, we show later (in Table 12 Panel B) that our results are robust to variation in this bound. Both the constraint on the regression's fitted values and the constraint on WLS observation weights bind in the sample we study.

The results of the first stage regression generating the state variable $EVAR_t$ are reported in Table 1 Panel A. Perhaps not surprisingly, past realized variance strongly predicts future realized variance. More importantly, the regression documents 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. In particular, the fact that we find that very persistent variables like PE and DEF forecast next period's volatility indicates a potential important role in volatility news for lower frequency or long-run movements in stochastic volatility.

We argue that the links we find 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 be associated with higher $RVAR$. The result that higher PE predicts *higher* $RVAR$ might 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. Since the default spread should also generally depend on the equity premium and since most of the variation in PE is due to variation in the equity premium, for a given value of the default spread, a relatively high value of PE implies a relatively higher level of future volatility. Thus PE cleans up the information in DEF concerning future volatility.

The R^2 of this regression is nearly 37%. The relatively low R^2 masks the fact that the fit is indeed quite good, as we can see from Figure 2, in which $RVAR$ and $EVAR$ are plotted together. The R^2 is heavily influenced by occasional spikes in realized variance, which the simple linear model we use is not able to capture. Indeed, our WLS approach downweights the importance of those spikes in the estimation procedure.

The online appendix to the paper reports descriptive statistics for these variables for the full sample, an early sample ending in 1963:3, and a modern sample beginning in 1963:4. Consistent with Campbell, Giglio and Polk (2012), we document high correlation between DEF and both PE and VS. The table also documents the persistence of both $RVAR$ and $EVAR$ (with autocorrelations of 0.52 and 0.74 respectively) and the high correlation between these variance measures and the default spread.

Perhaps the most notable difference between the two subsamples is that the correlation between PE and several of our other state variables changes dramatically. In the early sample, PE is quite negatively correlated with both $RVAR$ and VS. In the modern sample, PE is essentially uncorrelated with $RVAR$ and positively correlated with VS. As a consequence, since $EVAR$ is just a linear combination of our state variables, the correlation between PE and $EVAR$ changes sign across the two samples. In the early sample, this correlation is very negative, with a value of -0.51. This strong negative correlation reflects the high volatility that occurred during the Great Depression when prices were relatively low. In the modern sample, the correlation is positive at 0.14. The positive correlation simply reflects the economic fact that episodes with high volatility and high stock prices, such as the technology boom of the late 1990s, were more prevalent in this subperiod than episodes with high volatility and low stock prices, such as the recession of the early 1980s. In section 5.2 below, we discuss the implications of these correlation changes for the implied volatility beta of the aggregate stock index.

4.3 Estimation of the VAR and the news terms

We estimate a first-order VAR as in equation (15), where $\mathbf{x}_{t+1}$ is a 6×1 vector of state variables ordered as follows:

$$\mathbf{x}_{t+1} = [r_{M,t+1} \; EVAR_{t+1} \; PE_{t+1} \; TY_{t+1} \; DEF_{t+1} \; VS_{t+1}]$$

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 $\mathbf{\Gamma}$ 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 $\mathbf{\Sigma}$, 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 the previous section 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, where 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). 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. Of course, the relatively high degree of correlation among PE, DEF, VS, and $EVAR$ complicates the interpretation of the individual effect 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.[12]

Panel C of Table 1 reports the sample correlation and autocorrelation 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. There are a couple of aspects of these results to note. For one thing, a comparison of the standard deviations of the unscaled and scaled residuals provides a rough indication of the effectiveness of our empirical solution to the heteroskedasticity of the VAR. In general, the standard deviations of the scaled residuals are several times larger than their unscaled counterparts. More specifically, our approach implies that the scaled return residuals should have unit standard deviation. Our implementation results in a sample standard deviation of 1.14, that is relatively close to the model's predicted value of 1.

Additionally, a comparison of the unscaled and scaled autocorrelation matrices reveals that much of the sample autocorrelation in the unscaled residuals is eliminated by our WLS approach. For example, the unscaled residuals in the regression forecasting the log real return have an autocorrelation of -0.074. The corresponding autocorrelation of the scaled return residuals is essentially zero, 0.002. Though the scaled residuals in the $EVAR$, PE and DEF regression still display some negative autocorrelation, the unscaled residuals are much more negatively autocorrelated.

Table 2 reports the coefficients of a regression of the squared unscaled residuals $\sigma_t u_{t+1}$ 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 1.9, rather than the theoretically expected value of one, but this coefficient is sensitive to the weighting scheme used in the regression. The fact that $EVAR$ significantly predicts with a positive sign all the squared errors of the VAR shows that the volatilities of all innovations are driven by a common factor, as we assume, although of course it is possible that empirically, other factors also influence the volatilities of certain variables.

The top panel of Table 3 presents the variance-covariance matrix and the standard deviation/correlation matrix of the news terms, estimated as described above. Consistent with previous research, we find that discount-rate news is twice as volatile as cash-flow news.

The interesting new results in this table concern the variance news term N_V. First, news about future variance has significant volatility, with nearly a third of the variability of discount-rate news. Second, variance news is negatively correlated (-0.22) with cash-flow news: as one might expect from the literature on the "leverage effect" (Black 1976, Christie

[12] One worry is that many of the elements of the transition matrix are estimated imprecisely. Though these estimates may be zero, their non-zero but statistically insignificant in-sample point estimates, in conjunction with the highly-nonlinear function that generates discount-rate and volatility news, may result in misleading estimates of risk prices. However, Table 11 Panel B shows that results are qualitatively similar if we instead employ a partial VAR where, via a standard iterative process, only variables with t-statistics greater than 1.0 are included in each VAR regression.

1982), news about low cash flows is associated with news about higher future volatility. This finding makes it unappealing to assume that variance news and cash-flow news are uncorrelated, as would be required for the validity of the model solution in Bansal, Kiku, Shaliastovich, and Yaron (2013). Third, N_V correlates negatively (-0.09) with discount-rate news, indicating that news of high volatility tends to coincide with news of low future real returns.[13] The net effect of these correlations, documented in the lower left panel of Table 3, is a slightly negative correlation of -0.02 between our measure of volatility news and contemporaneous market returns (for related research see French, Schwert, and Stambaugh 1987).

The lower right panel of Table 3 reports the decomposition of the vector of innovations $\sigma_t^2 u_{t+1}$ 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, TY, 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, we find that innovations in PE, DEF, and VS are associated with news of higher future volatility. This panel also indicates that all state variables with the exception of TY are statistically significant in terms of their contribution to at least one of the three news terms. We choose to leave TY in the VAR, though its presence in the system makes little difference to our conclusions.

Figure 3 plots the smoothed series for N_{CF}, $-N_{DR}$ and N_V using an exponentially-weighted moving average with a quarterly decay parameter of 0.08. This decay parameter implies a half-life of six years. The pattern of N_{CF} and $-N_{DR}$ we find is consistent with previous research. As a consequence, we focus on the smoothed series for market variance news. There is considerable time variation in N_V, and in particular 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 3 to 5 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 crash of the stock market. 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 instead characterized by strongly negative cash-flow news, together with a spike in volatility of the highest magnitude in our sample. The recovery from the financial crisis has brought positive cash-flow news together with news about lower future volatility.

[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.

4.4 Predicting long-run volatility

The predictability of volatility, and especially of its long-run component, is central to this paper. In the previous sections, we have shown that volatility is strongly predictable, and it is predictable in particular 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 like PE and DEF.

We now ask how well our proposed long-run volatility forecast captures the long-horizon component of volatility. In Table 4 we regress realized, discounted, annualized long-run variance up to period h,

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

on both our VAR forecast and some alternative forecasts of long-run variance.[14] We focus on the 10-year horizon ($h = 40$) as longer horizons come at the cost of fewer independent observations; however, Table 3 in the online appendix confirms that our results are robust to horizons ranging from one to 15 years.

GARCH models

We estimate two standard GARCH-type models, specifically designed to capture the long-run component of volatility. The first one is the two-component EGARCH model proposed by Adrian and Rosenberg (2008). This model assumes the existence of two separate components of volatility, one of which is more persistent than the other, and therefore will tend to capture the long-run dynamics of the volatility process. The other model we estimate is the FIGARCH model of Baillie, Bollerslev, and Mikkelsen (1996), in which the process for volatility is modeled as a fractionally-integrated process, and whose slow, hyperbolic rate of decay of lagged, squared innovations potentially captures long-run movements in volatility better. We first estimate both GARCH models using the full sample of daily returns and then generate the appropriate forecast of $LHRVAR_{40}$.[15] To these two models, we add the set of variables from our VAR, and compare the forecasting ability of these different models.

Table 4 Panel A reports the results of forecasting regressions of long run volatility $LHRVAR_{40}$ using different specifications. The first regression presents results using the state variables in our VAR, each included separately. The second regression predicts $LHRVAR_{40}$

[14]Note that we measure $LHRVAR$ in annual units. In particular, we rescale by the sum of the weights ρ^j to maintain the scale of the coefficients in the predictive regressions across different horizons.

[15]We start our forecasting exercise in January 1930 so that we have a long enough history of past returns to feed the FIGARCH model. Other long-run GARCH models could be estimated in a similar manner, for example the FIEGARCH model of Bollerslev and Mikkelsen (1996).

with the horizon-specific forecast implied by our VAR (VAR_{40}). The third and fourth regressions forecast $LHRVAR_{40}$ with the corresponding forecast from the EGARCH model (EG_{40}) and the FIGARCH model (FIG_{40}) respectively. The fifth and sixth regressions join the VAR variables with the two GARCH-based forecasts, one at a time. The seventh and eighth regressions conduct a horse race between VAR_{40} and FIG_{40} and between VAR_{40} and DEF. Regressions nine through 13 focus on the forecasting ability of our two key state variables, DEF and PE; we discuss these specifications in more detail below.

First note that both the EGARCH and FIGARCH forecasts by themselves capture a significant portion of the variation in long-run realized volatility: both have significant coefficients, and both have nontrivial R^2s. Our VAR variables provide as good or better explanatory power, and $RVAR$, PE and DEF are strongly statistically significant. Online Appendix Table 3 documents that these conclusions are true at all horizons (with the exception of $RVAR$ at $h = 20$, i.e. 5 years). Finally, the coefficient on the VAR-implied forecast, VAR_{40}, is 0.989. This estimate is not only significantly different from zero but also not significantly different from one. This finding indicates that our VAR is able to produce forecasts of volatility that not only go in the right direction, but are also of the right magnitude, even at the 10-year horizon.

Very interesting results appear once we join our variables to the two GARCH models. Even after controlling for the GARCH-based forecasts (which render $RVAR$ insignificant), PE and DEF significantly predict long-horizon volatility, and the addition of the VAR state variables strongly increases the R^2. We further show that when using the VAR-implied forecast together with the FIGARCH forecast, the coefficient on VAR_{40} is still very close to one and always statistically significant while the FIGARCH coefficient moves closer to zero (though it remains statistically significant at the 10-year horizon).

We develop an additional test of our VAR-based model of stochastic volatility from the idea that the variables that form the VAR – in particular the strongest of them, DEF – should predict volatility at long horizons only *through* the VAR, not *in addition* to it. In other words, the VAR forecasts should ideally represent the best way to combine the information contained in the state variables concerning long-run volatility. If true, after controlling for the VAR-implied forecast, DEF or other variables that enter the VAR should not significantly predict future long-run volatility. We test this hypothesis by running a regression using *both* the VAR-implied forecast and DEF as right-hand side variables. We find that the coefficient on VAR_{40} is still not significantly different from one, while the coefficient on DEF is small and statistically indistinguishable from zero. The online appendix shows that this finding is true at all horizons we consider.

Interpretation of the long-run VAR forecast

The bottom part of Table 4 Panel A examines more carefully the link between DEF and $LHRVAR_{40}$. Regressions 9 through 13 in the table forecast $LHRVAR_{40}$ with PE, DEF, PEO (PE orthogonalized to DEF), and $DEFO$ (DEF orthogonalized to PE). These regressions show that by itself, PE has no information about low-frequency variation in volatility. In contrast, DEF forecasts nearly 22% of the variation in $LHRVAR_{40}$. And once DEF is orthogonalized to PE, the R^2 increases to 51%. Adding PEO has little effect on the R^2. We argue that this is clear evidence of the strong predictive power of the orthogonalized component of the default spread.

Recall our simple interpretation of these results. DEF contains information about future volatility as risky bonds are short the option to default. However, DEF also contains information about future aggregate risk premia. We know from previous work that most of the variation in PE is about aggregate risk premia. Therefore, including PE in the volatility forecasting regression cleans up variation in DEF due to aggregate risk premia and thus sharpens the link between DEF and future volatility. Since 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.

In Figure 4, we provide a visual representation of the volatility-forecasting power of our key VAR state variables and our interpretation of the results. The top panel 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 how both variables track long-run movements in long run volatility. To isolate the contribution of the default spread in predicting long run volatility, the bottom panel plots $LHRVAR_{40}$ together with $DEFO$. In general, the improvement in fit moving from the top panel to the bottom panel is clear.

More specifically, 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 expectations of future market volatility. However, once the declining equity premium over that period is taken into account (as shown by the rapid increase in PE), one recognizes that a PE-adjusted default spread in the late 1990s actually forecasted much higher volatility ahead.

Option-implied variance forecasts

As a further check on the usefulness of our VAR approach, we compare our variance forecasts to option-implied variance forecasts. Specifically, we use 1998–2011 prices of variance swaps ranging from one month to one year in maturity to compare the forecast of long-horizon variance at horizon h from our baseline VAR (VAR_h) to the corresponding VIX^2 at horizon h (VIX_h^2).[16] Since our VAR is quarterly, we study forecasts at the three-month, six-month, nine-month, and twelve-month horizons. The top panels of Figure 5 plot the time series of these forecasts for the three-month and twelve-month horizons. We find that forecasts from the two quite different methods line up well, though the VIX^2 forecasts are generally higher, especially near the end of the sample. The figure also shows that the VAR_h forecasts become smoother when the horizon is extended, relative to both the shorter-horizon VAR_h forecasts as well as the VIX_h^2 forecasts at the same horizon. Table 4 Panel B confirms these facts by reporting the mean, standard deviation, and correlation of these forecasts, along with the value for realized variance ($LHRVAR_h$) over the corresponding horizon. The VIX^2 forecasts are on average approximately 20% larger than their realized variance counterparts.

Table 4 Panel C reports regressions forecasting $LHRVAR_h$ using the VAR_h forecast, the VIX_h^2 forecast, or both together, at each horizon. Both the VAR and the option-based forecasts are individually statistically significant, though the coefficient on VAR_h is always closer to the predicted value of 1.0 at all horizons except for three months. The bottom panels of Figure 5 plot $LHRVAR_h$ against the fitted value from the VAR_h forecast and against the fitted value of the VIX_h^2 forecast for the three-month and twelve-month horizons. The figure confirms that VAR_h is as informative as VIX_h^2, if not more so. Indeed, Table 4 Panel C shows that when both forecasts are included in the regression, VAR_h subsumes VIX_h^2, remaining statistically and economically significant.

Taken together, the results in Table 1 Panel A and Table 4 make a strong case that credit spreads and valuation ratios contain information about future volatility not captured by simple univariate models, even those like the FIGARCH model or the two-component EGARCH model that are designed to fit long-run movements in volatility, and that our VAR method for calculating long-horizon forecasts preserves this information.

[16] As the VIX_h^2 measures do not discount future volatility, for this portion of the analysis, we do not discount either expectations of future variance when constructing our VAR_h measures or their realized variance counterparts when constructing $LHRVAR_h$.

5 Test Assets and Beta Measurement

5.1 Test assets

In addition to the six VAR state variables, our analysis also requires returns on a cross section of test assets. We construct several sets of portfolios for this purpose.

Characteristic-sorted test assets

Our primary cross section consists of the excess returns on the 25 ME- and BE/ME-sorted portfolios, studied in Fama and French (1993), extended in Davis, Fama, and French (2000), and made available by Professor Kenneth French on his web site.[17]

We consider two main subsamples: early (1931:3-1963:3) and modern (1963:4-2011:4) due to the findings in Campbell and Vuolteenaho (2004) of dramatic differences in the risks of these portfolios between the early and modern period. The first subsample is shorter than that in Campbell and Vuolteenaho (2004) as we require each of the 25 portfolios to have at least one stock as of the time of formation in June.

Risk-sorted test assets

Daniel and Titman (1997, 2012) and Lewellen, Nagel, and Shanken (2010) point out that it can be misleading to test asset pricing models using only portfolios sorted by characteristics known to be related to average returns, such as size and value. In particular, characteristic-sorted portfolios are likely to show some spread in betas identified as risk by almost any asset pricing model, at least in sample. When the model is estimated, a high premium per unit of beta will fit the large variation in average returns. Thus, at least when premia are not constrained by theory, an asset pricing model may spuriously explain the average returns to characteristic-sorted portfolios.

Our model has tightly constrained risk prices which protects us against this critique. Nonetheless, to alleviate this concern, we follow the advice of Daniel and Titman (1997, 2012) and Lewellen, Nagel, and Shanken (2010) and construct a second set of six portfolios double-sorted on past risk loadings to market and variance risk. First, we run a loading-estimation regression for each stock in the CRSP database where $r_{i,t}$ is the log stock return on stock i for month t.

$$\sum_{j=1}^{3} r_{i,t+j} = b_0 + b_{r_M} \sum_{j=1}^{3} r_{M,t+j} + b_{\Delta VAR} \sum_{j=1}^{3} \Delta VAR_{t+j} + \varepsilon_{i,t+3} \tag{27}$$

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

We calculate ΔVAR as a weighted sum of changes in the VAR state variables. The weight on each change is the corresponding value in the linear combination of VAR shocks that defines news about market variance. We choose to work with changes rather than shocks as this allows us to generate pre-formation loading estimates at a frequency that is different from our VAR. Namely, though we estimate our VAR using calendar-quarter-end data, our approach allows a stock's loading estimates to be updated at each interim month.

The regression is reestimated from a rolling 36-month window of overlapping observations for each stock at the end of each month. Since these regressions are estimated from stock-level instead of portfolio-level data, we use quarterly data to minimize the impact of infrequent trading. With loading estimates in hand, each month we perform a two-dimensional sequential sort on market beta and ΔVAR beta. First, we form three groups by sorting stocks on $\widehat{b}_{r_M}$. Then, we further sort stocks in each group to three portfolios on $\widehat{b}_{\Delta VAR}$ and record returns on these nine value-weight portfolios. The final set of risk-sorted portfolios are the two sets of three $\widehat{b}_{r_M}$ portfolios within the extreme $\widehat{b}_{\Delta VAR}$ groups. To ensure that the average returns on these portfolio strategies are not influenced by various market-microstructure issues plaguing the smallest stocks, we exclude the five percent of stocks with the lowest ME from each cross-section and lag the estimated risk loadings by a month in our sorts.

Characteristic- and risk-sorted test assets

We also consider equity portfolios that are formed based on both characteristics and past risk loadings. 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, which increase in value when volatility increases, all else equal. 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: BE/ME and idiosyncratic volatility (*ivol*).

We first sort stocks into tritiles based on BE/ME and then into tritiles based on *ivol*. We follow Ang, Hodrick, Xing, and Zhang (2006) and others and estimate *ivol* as the volatility of the residuals from a Fama and French (1993) three-factor regression using daily returns within each month. Finally, we split each of these nine portfolios into two subsets based on pre-formation estimates of *simple* volatility beta, $\widehat{\beta}_{\Delta VAR}$, estimated as above but in a simple regression that does not control for the market return. One might expect that sorts on simple rather than partial betas will be more effective in establishing a link between pre-formation and post-formation estimate of volatility beta, since the market is correlated with volatility news. As before, we exclude the bottom five percent of stocks based on market capitalization and lag our loadings and idiosyncratic volatility estimates by one month.

Non-equity test assets

We consider several sets of non-equity test assets. We generate a parsimonious cross section of option, bond, and equity returns for the 1986:1–2011:4 time period based on the findings in Fama and French (1993) and Coval and Shumway (2001). In particular, we use the S&P 100 index straddle returns studied by Coval and Shumway.[18] We also include proxies for the two components of the risky bond factor of Fama and French (1993) which we measure using the return on the Barclays Capital High Yield Bond Index ($HYRET$) and the return on Barclays Capital Investment Grade Bond Index ($IGRET$). When pricing the straddle and risky bond return series, we include the returns on the market ($RMRF$), size (SMB), and value (HML) equity factors of Fama and French (1993) as they argue these factors do a good job describing the cross section of average equity returns.

We also study the cross-section of currency portfolios, where developed-country currencies have been dynamically allocated to portfolios based on their interest rates as in Lustig, Roussanov, and Verdelhan (2011).[19] The currency portfolios cover the period 1984:1–2010:1.

Finally, from the options market, we generate monthly returns on 12 synthetic variance forward contracts, covering the period 1998:4–2011:4. We construct these returns as follows. First, we construct a panel of implied variance swap prices using option data from OptionMetrics, for maturities n ranging from 1 to 12 months: $VIX^2_{n,t}$. Under the assumption that returns follow a diffusion, we will 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 CBOE to construct the 30-day VIX, applying it to maturities up to 12 months. We then compute synthetic variance forward prices as: $F_{n,t} = VIX^2_{n,t} - VIX^2_{n-1,t}$. These forwards allow us to isolate claims to variance at a specific horizon n (focusing on the variance realized between $n-1$ and n). The returns to these forwards are computed as $R_{n,t} = \frac{F_{n-1,t}}{F_{n,t-1}} - 1$, where $F_{0,t} = RVAR_t$. These quite volatile portfolios are rescaled by a factor of 20 so that their return volatility is broadly similar in magnitude to our other test assets and are then aggregated to a quarterly frequency.

5.2 Beta measurement

We now examine the validity of an unconditional version of the first-order condition in equation (19). We modify equation (19) in three ways. First, we use simple expected returns on the left-hand side to make our results easier to compare with previous empirical studies. Second, we condition down equation (19) to avoid having to estimate all required conditional moments. Finally, we cosmetically multiply and divide all three covariances by the sample variance of the unexpected log real return on the market portfolio. By doing so,

[18] Specifically, the series we study includes only those straddle positions where the difference between the options' strike price and the underlying price is between 0 and 5. We thank Josh Coval and Tyler Shumway for providing their updated data series to us.

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

we can express our pricing equation in terms of betas, facilitating comparison to previous research. These modifications result in the following asset-pricing equation

$$E[R_i - R_f] = \gamma\sigma_M^2\beta_{i,CF_M} + \sigma_M^2\beta_{i,DR_M} - \frac{1}{2}\omega\sigma_M^2\beta_{i,V_M}, \tag{28}$$

where

$$\begin{aligned} \beta_{i,CF_M} &\equiv \frac{Cov(r_{i,t}, N_{CF,t})}{Var(r_{M,t} - E_{t-1}r_{M,t})}, \\ \beta_{i,DR_M} &\equiv \frac{Cov(r_{i,t}, -N_{DR,t})}{Var(r_{M,t} - E_{t-1}r_{M,t})}, \\ \text{and } \beta_{i,V_M} &\equiv \frac{Cov(r_{i,t}, N_{V,t})}{Var(r_{M,t} - E_{t-1}r_{M,t})}. \end{aligned}$$

We price the average excess returns on our test assets using the unconditional first-order condition in equation (28) and the quadratic relationship between the parameters ω and γ given by equation (24). As a first step, 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 the previous section. Specifically, 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.

Characteristic-sorted betas

Table 5 Panel A shows the estimated betas for the 25 size- and book-to-market portfolios over the 1931-1963 period. 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, while 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 have both higher cash-flow and higher discount-rate betas than growth stocks. An equal-weighted average of the extreme value stocks across size quintiles has a cash-flow beta 0.12 higher than an equal-weighted average of the extreme growth stocks. The difference in estimated discount-rate betas, 0.20, is in the same direction. Similar to value stocks, small stocks have higher cash-flow betas and discount-rate betas than large stocks in this sample (by 0.14 and 0.34, respectively, for an equal-weighted average of the smallest stocks across value quintiles relative to an equal-weighted average of the largest stocks). These differences are extremely similar to those in Campbell and Vuolteenaho (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 5 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 size quintiles has a volatility beta 0.05 lower than an equal-weighted average of the extreme growth stocks. Similarly, an equal-weighted average of the smallest stocks across value quintiles has a volatility beta that is 0.04 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 6 Panel A reports the corresponding estimates for the post-1963 period. As documented in this subsample by Campbell and Vuolteenaho (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 0.13 lower than the volatility beta of an equal-weighted average of the extreme growth stocks, a difference that is more than 42% higher than the corresponding difference in the early period.[20]

These results imply that in the post-1963 period where the CAPM has difficulty explaining the low returns on growth stocks relative to value stocks, growth stocks are relative hedges for two key aspects of the investment opportunity set. Consistent with Campbell and Vuolteenaho (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.

[20] Our findings are in sharp contrast to Bansal, Kiku, Shaliastovich, and Yaron (2013), who find that value-minus-growth portfolios are volatility hedges. See their Table VII where they estimate that the β_V of a portfolio that is long the top quintile of value stocks and short the corresponding bottom quintile is 2.8, roughly 60% of the absolute value of the market's β_V; see their Table X where those numbers are 1.0 and 81% respectively; and see their Table XI where the numbers are 1.1 and 62% respectively. Their finding that value-minus-growth bets are volatility hedges is hard to reconcile with theory (real option models such as McQuade 2012) and stylized facts (notably the performance of value-minus-growth bets during the Great Depression, the Tech Boom, and the Great Recession).

The changing volatility beta of the market portfolio

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 -0.06 while over the 1964-2011 period this average becomes 0.09. Of course, given the strong positive link between PE and volatility news documented in the lower right panel of Table 3, one should not be surprised that the market's β_V can be positive. Moreover, given the change in sign over time in PE's correlation with some of the key state variables driving $EVAR$ documented in the online appendix, one should not be surprised that β_V changes sign as well. Nevertheless, we study this change in sign more carefully.

Figure 6 shows scatter plots with the early period as blue triangles and the modern period data as red asterisks. The top two plots in this figure emphasize that variance news betas are not the same as $RVAR$ betas. The top left portion of the figure plots the market return against $RVAR$. This plot shows that the market does do poorly when realized variance is high, and that this is the case in both subsamples. In fact, this relation is slightly more negative in the modern period. However, our theory tells us that long-horizon investors care about low frequency movements in volatility. The top right portion of the figure plots the market return against volatility news, N_V. Consistent with the estimates in Tables 5 and 6 in the paper, the relation between the market return and N_V is negative in the early period and positive in the modern period.[21] This plot shows that the estimates are robust and not driven by outliers.

The bottom two plots in this figure illustrate what drives this relation in our VAR. The bottom left of the figure plots PE against $DEFO$, our simple proxy for news about long-horizon variance. It is easy to see that the market's PE is high when $DEFO$ is low in the early period, but this relation reverses in the latter period. The bottom right of the figure plots market returns against the contemporaneous change in $DEFO$, showing a negative relation in the early period and a positive relation in the modern period. In other words, the orthogonalization of DEF to PE that creates $DEFO$ is valid over the whole sample, but conceals negative comovement in the early period and positive comovement in the modern period.

In summary, Figure 5 highlights the important distinction between single-period realized variance $RVAR$ and long-run volatility news, and confirms that the sign change in the market's volatility beta from the early to the modern period can be seen in simple plots of the market return against the change in our key state variable, the PE-adjusted default spread. Table 11 examines the robustness of this finding to different VAR specifications and estimation methods.

[21] Straddle returns are negatively correlated with the return on the market portfolio in the 1986:1-2011:4 sample. This negative correlation is not inconsistent with the positive correlation we find between the market return and N_V in the modern sample as the straddle portfolio consists of one-month maturity options and thus should respond to short-term volatility expectations.

Risk-sorted betas

Table 5 Panel B shows the estimated betas for the six risk-sorted portfolios over the 1931-1963 period. The portfolios are organized in a rectangular matrix with low market-beta stocks at the left, high market-beta stocks at the right, low volatility-beta stocks at the top, and high volatility-beta stocks at the bottom. At the right edge of the matrix we report the differences between the high market-beta and the low market-beta portfolios in each volatility-beta group; along the bottom of the matrix we report the differences between the high volatility-beta and the low volatility-beta portfolios in each market-beta category. As in Panel A, the top matrix displays post-formation cash-flow betas, the middle matrix displays post-formation discount-rate betas, while the bottom matrix displays post-formation volatility betas.

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. An equal-weighted average of the high market-beta stocks across the two volatility-beta categories has a cash-flow beta 0.19 higher, and a discount-rate beta 0.44 higher, than an equal-weighted average of the low market-beta stocks. Similar to high market-beta stocks, low volatility-beta stocks have higher cash-flow betas and discount-rate betas than high volatility-beta stocks in this subsample (by 0.06 and 0.11, respectively, for an equal-weighted average of the low volatility-beta stocks across the three market-beta categories relative to a corresponding equal-weighted average of the high volatility-beta stocks).

High market-beta stocks and low volatility-beta stocks are also riskier in terms of volatility betas. An equal-weighted average of the high market-beta stocks across volatility-beta categories has a post-formation volatility beta 0.04 lower than an equal-weighted average of the low market-beta stocks. Similarly, an equal-weighted average of the low volatility-beta stocks across market-beta categories has a post-formation volatility beta that is 0.02 lower than an equal-weighted average of the high volatility-beta stocks. In summary, high market-beta and low volatility-beta stocks were unambiguously riskier than low market-beta and high volatility-beta stocks over the 1931-1963 period.

Table 6 Panel B shows the estimated betas for the six risk-sorted portfolios over the modern (post-1963) period. In the modern period, high market-beta stocks again have higher cash-flow and higher discount-rate betas than low market-beta stocks, by 0.08 and 0.55 respectively. However, high market-beta stocks are no longer riskier in terms of volatility betas. Now, an equal-weighted average of the high market-beta stocks across the two volatility-beta categories has a post-formation variance beta 0.07 higher than a corresponding equal-weighted average of the low market-beta stocks. Since, in the three-beta model, covariation with aggregate volatility has a negative premium, the three-beta model can potentially explain why stocks with high past market betas have offered relatively little extra return in the modern period.

In the modern period, sorts on volatility beta continue to generate an economically and statistically significant spread in post-formation volatility beta. An equal-weighted average of low volatility-beta stocks across the three market-beta categories has a post-formation volatility beta that is 0.06 lower than the post-formation volatility beta of a corresponding equal-weighted average of high volatility-beta stocks. Sorts on volatility beta also generate spread in discount-rate betas, but essentially no spread in cash-flow betas in the post-1963 period.

Characteristic- and risk-sorted betas

We also examine test assets that are formed based on both characteristics *and* risk estimates. Table 6 Panel C shows the estimated betas for the 18 BE/ME-*ivol*-$\widehat{\beta}_{\Delta VAR}$-sorted portfolios over the post-1963 period. Consistent with the above results, stocks with lower BE/ME have higher post-formation volatility betas. The composite portfolio, denoted $P1$, that is long the equal-weight average of the value portfolios and is short the equal-weight average of the growth portfolios has a volatility beta of -0.10 that is statistically significant. Table 2 of the Appendix shows that $P1$ has an average return of 9.28%/year.

Among stocks with low BE/ME, firms with higher *ivol* have higher post-formation volatility betas. The volatility beta of the composite portfolio, denoted $P2$ Growth, that is long the high idiosyncratic portfolio and short the low idiosyncratic portfolio among growth stocks is 0.06 and has an average return of -6.41%/year. Among stocks with high BE/ME, firms with higher *ivol* have only marginally higher post-formation volatility betas (the corresponding portfolio $P2$ Value has a volatility beta of only 0.02). Interestingly, among stocks with high BE/ME, firms with higher *ivol* have significantly *higher* average returns ($P2$ Value has an average return of 5.97%/year).

We argue 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.

These results have the potential to explain the puzzling finding that high idiosyncratic-volatility stocks have lower average returns than low idiosyncratic-volatility stocks, as well as the fact that the unconditional *ivol* effect is non-monotonic (Ang, Hodrick, Xing, and Zhang 2006). In particular, the relation between *ivol* and average returns depends on a firm's growth prospects, proxied by BE/ME. Our analysis shows that the reason that *ivol* is positively related to average returns among value stocks is that the relation between *ivol* and β_{CF} is much stronger among value stocks, generating a 50% larger spread in β_{CF}, while the relation between *ivol* and β_V is much weaker. These results may also explain why the *ivol* effect appears to be less robust in some samples using different methodologies (Bali and Cakici 2008) and even switches sign in others (Fu 2009). Certainly some weighting schemes and samples may lean towards high idiosyncratic-volatility stocks that are relatively more value than growth.

Finally, we show that within each of the nine BE/ME-*ivol*-sorted portfolios, our sort on pre-formation simple volatility beta continues to generate a reasonable amount of spread in post-formation volatility beta and average returns. The volatility beta of the composite portfolio, denoted $P3$, that is long the equal-weight average of the high $\widehat{\beta}_{\Delta VAR}$ portfolios and is short the equal-weight average of the low $\widehat{\beta}_{\Delta VAR}$ portfolios is 0.07 and highly statistically significant. This portfolio has an average return of -0.85%/year. Taken together, the findings from the characteristic- and risk-sorted test assets suggest that volatility betas vary with multiple stock characteristics, and that techniques that take this into account may be more effective in generating a spread in post-formation volatility beta.

Non-equity betas

Table 6 Panel D reports the three ICAPM betas of the S&P 100 index straddle position analyzed in Coval and Shumway (2001) along with the corresponding ICAPM betas of the three equity factors and the default bond factor of Fama and French (1993) over the period 1986:1 - 2011:4. Consistent with the nature of a straddle bet, we find that the straddle has a very large volatility beta of 0.38. It also has a large negative discount-rate beta of -1.71 and a large (relatively speaking) negative cash-flow beta of -0.39. As one would expect, the betas of the Fama-French equity factors are consistent with the findings for the size- and book-to-market-sorted portfolios in Table 6 Panel B. Finally, the riskier component of Fama and French's (1993) risky bond factor, $HYRET$, has a cash-flow beta of 0.06, a discount-rate beta of 0.26, and a volatility beta of -0.05. These betas are economically and statistically significant, unlike those of the safer component, $IGRET$. The difference in volatility beta between $HYRET$ and $IGRET$ is consistent with the fact that risky corporate debt is short the option to default.

Table 6 Panel E reports the three ICAPM betas of the five interest-rate-sorted developed-country currency portfolios from Lustig, Roussanov, and Verdelhan (2011). High-interest-rate countries have higher cash-flow and discount-rate betas and lower volatility betas than their low-interest-rate counterparts. Thus, high-interest-rate currencies are unambiguously riskier over the 1984:1-2010:1 period.

Table 6 Panel F reports the ICAPM betas on the variance forward positions. The one-month variance forward has a positive variance beta which is intuitive. Interestingly, variance betas decline as the horizon increases, though the results are not statistically significant. The one-month variance forward has cash-flow and discount-rate betas that are both negative and statistically significant. These betas increase as the horizon increases, becoming effectively zero at the one-year horizon. These changes are broadly consistent with the term structure of average returns on these variance forwards, reported in the online appendix.

6 Pricing the Cross-Section of Security Returns

We next turn to pricing the cross section with these three ICAPM betas. We evaluate the performance of five asset-pricing models: 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 Campbell and Vuolteenaho (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 equation (24), 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.

Each model is estimated in two different forms: one with a restricted zero-beta rate equal to the Treasury-bill rate as in the Sharpe-Lintner version of the CAPM, and one with an unrestricted zero-beta rate following Black (1972). Allowing for an unrestricted zero-beta rate may be particularly important given the extensive evidence in Krishnamurthy and Vissing-Jørgensen (2012) that Treasury Bills provide convenience benefits in terms of liquidity and safety.[22]

6.1 Characteristic-sorted test assets

Table 7 reports results for the early sample period 1931-1963, using 25 size- and book-to-market-sorted portfolios as test assets. The table has ten columns, two specifications for each of our five asset pricing models. The first 16 rows of Table 7 are divided into four sets of four rows. The first set of four rows corresponds to the zero-beta rate (in excess of the Treasury-bill rate), the second set to the premium on cash-flow beta, the third set to the premium on discount-rate beta, and the fourth 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 annualizes this estimate, multiplying by 400 to aid in interpretation. These parameters are estimated from a cross-sectional regression

$$\overline{R}_i^e = g_0 + g_1\widehat{\beta}_{i,CF_M} + g_2\widehat{\beta}_{i,DR_M} + g_3\widehat{\beta}_{i,V_M} + e_i, \tag{29}$$

where a bar denotes time-series mean and $\overline{R}_i^e \equiv \overline{R}_i - \overline{R}_{rf}$ denotes the sample average simple excess return on asset i. The third and fourth rows present two alternative standard errors of the monthly estimate, described below.

[22] Krishnamurthy and Vissing-Jorgensen (2012) conclude, "Our finding of a convenience demand for Treasury debt suggests caution against the common practice of identifying the Treasury interest rate with models' riskless interest rate." Similar arguments can be found in Duffie and Singleton (1997) and Hull, Predescu, and White (2004).

Below the premia estimates, we report the R^2 statistic for a cross-sectional regression of average returns on our test assets onto the fitted values from the model. We also report a composite pricing error, computed as a quadratic form of the pricing errors. The weighting matrix in the quadratic form is a diagonal matrix with the inverse of the sample test asset return volatilities on the main diagonal.

Standard errors are produced with a bootstrap from 10,000 simulated realizations. Our bootstrap experiment samples test-asset returns and first-stage VAR errors, and uses the first-stage and second-stage WLS VAR estimates in Table 1 to generate the state-variable data.[23] We partition the VAR errors and test-asset returns into two groups, one for 1931 to 1963 and another for 1963 to 2011, which enables us to use the same simulated realizations in subperiod analyses. The first set of standard errors (labeled A) conditions on estimated news terms and generates betas and return premia separately for each simulated realization, while the second set (labeled B) also estimates the first-stage and second-stage VAR and the news terms separately for each simulated realization. Standard errors B thus incorporate the considerable additional sampling uncertainty due to the fact that the news terms as well as betas are generated regressors.

Two alternative 5-percent critical values for the composite pricing error are produced with a bootstrap method similar to the one we have described above, except that the test-asset returns are adjusted to be consistent with the pricing model before the random samples are generated. Critical values A condition on estimated news terms, while critical values B take account of the fact that news terms must be estimated.

Finally, Table 7 reports 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 three-beta ICAPM estimates are constrained so that both γ and the implied ω are strictly positive.

Results for the early period

Table 7 shows that in the 1931-1963 period, all our models explain the cross-section of stock returns reasonably well. The cross-sectional R^2 statistics are almost 56% for both forms of our three-beta ICAPM. Both the Sharpe-Lintner and Black versions of the CAPM do a slightly poorer job describing the cross section (both R^2 statistics are roughly 53%). The two-beta ICAPM of Campbell and Vuolteenaho (2004) performs slightly better than the CAPM and slightly worse than the three-beta ICAPM. None of the theoretically-motivated models considered are rejected by the data based on the composite pricing test. Consistent with the claim that the three-beta model does a good job describing the cross-section, Table 7 shows that the constrained and the unrestricted factor model barely improve pricing relative to the three-beta ICAPM.

[23] When simulating the bootstrap, we drop realizations which would result in negative $RVAR$ and redraw.

We can quantify the role that volatility betas play in our model. For the Black version of the three-beta ICAPM, the spread in volatility betas across the 25 size- and book-to-market-sorted portfolios generates an annualized spread in average returns of 1.6% compared to a comparable spread of 7.3% and 3.2% for cash-flow and discount-rate betas. Variation in volatility betas accounts for 2% of the variation in explained returns compared to 38% and 7% for cash-flow and discount-rate betas respectively. The remaining 53% of the explained variation in average returns is due to the covariation among the three types of betas.

Figure 7 provides a visual summary of the early-period results. The figure plots the predicted average excess return on the horizontal axis and the actual sample average excess return on the vertical axis, for test asset returns measured relative to the Treasury bill rate. Results are shown for the Sharpe-Lintner and Black versions of the CAPM and the three-beta ICAPM. The difficulty in distinguishing the models is visually apparent for this sample period.

Results for the modern period

Results are very different in the 1963-2011 period. Table 8 shows that in this period, both versions of the CAPM do a very poor job of explaining cross-sectional variation in average returns on portfolios sorted by size and book-to-market. When the zero-beta rate is left as a free parameter, the cross-sectional regression picks a negative premium for the CAPM beta and implies an R^2 of roughly 5%. When the zero-beta rate is constrained to the risk-free rate, the CAPM R^2 falls to roughly -37%. Both versions of the static CAPM are easily rejected at the five-percent level by both sets of critical values.

In the modern period, the unconstrained zero-beta rate version of the two-beta Campbell and Vuolteenaho (2004) model does a better job describing the cross section of average returns than the CAPM. However, the implied coefficient of risk aversion, 20.7, is arguably extreme.

The three-beta model with the restricted zero-beta rate also does a poor job explaining cross-sectional variation in average returns across our test assets. However, if we continue to restrict the risk price for discount-rate and variance news but allow an unrestricted zero-beta rate, the explained variation increases to roughly 63%, almost three-quarters larger than the R^2 of the corresponding two-beta ICAPM. The estimated risk price for cash-flow beta is an economically reasonable 21.5 percent per year with an implied coefficient of relative risk aversion of 6.9 (equal to the theoretical maximum consistent with the existence of a model solution). Neither version of our intertemporal CAPM with stochastic volatility is rejected at the 5-percent level by either set of critical values.

Once again, we can quantify the role that volatility betas play in our model. For the Black version of the three-beta ICAPM, the spread in volatility betas across the 25 size- and book-to-market-sorted portfolios generates an annualized spread in average returns of

5.2% compared to a comparable spread of 2.8% and 2.2% for cash-flow and discount-rate betas. Variation in volatility betas accounts for 104% of the variation in explained returns compared to 19% for cash-flow betas as well as 13% for discount-rate betas. Covariation among the three types of betas is responsible for the remaining -36% of explained variation in average returns.

Figure 8 provides a visual summary of the modern-period results. The poor performance of the CAPM in this sample period is immediately apparent. The version of the ICAPM with a restricted zero-beta rate, equal to the risk-free rate or Treasury bill rate, generates some cross-sectional spread in predicted returns that lines up qualitatively with average realized returns. However, almost all returns are underpredicted because stocks are estimated to be volatility hedges in the modern period, so the model implies a relatively low equity premium. This problem disappears when we free up the zero-beta rate in the ICAPM, adding the spread between the zero-beta rate and the Treasury bill rate to the predicted excess return over the bill rate.

The relatively poor performance of the risk-free rate version of the three-beta ICAPM is due to the derived link between γ and ω. To show this, Figure 8 provides two contour plots (one each for the risk-free and zero-beta rate versions of the model in the top and bottom panels of the figure respectively) of the R^2 resulting from combinations of (γ,ω) ranging from (0,0) to (40,40). On the same figure we also plot the relation between γ and ω derived in equation (24). The top panel of Figure 9 shows that even with the intercept restricted to zero, R^2's are as high as 70% for some combinations of (γ,ω). Unfortunately, as the plot shows, these combinations do not coincide with the curve implied by equation (24). Once the zero-beta rate is unconstrained, the contours for R^2's greater than 60% cover a much larger area of the plot and coincide nicely with the ICAPM relation of equation (24).

Consistent with the contour plots of Figure 9, the pricing results in Table 8 based on the partially-constrained factor model further confirms that the link between γ and ω is responsible for the poor fit of the restricted zero-beta rate version of the three-beta ICAPM in the modern period. When removing the constraint linking γ and ω but leaving the constraint on the discount-rate beta premium in place, the R^2 increases from -109% to 74%. Moreover, the risk prices for γ and ω remain economically large and of the right sign.

6.2 Alternative test assets

Risk-sorted test assets

We confirm that the success of the three-beta ICAPM is robust by expanding the set of test portfolios beyond the 25 size- and book-to-market-sorted portfolios. Our focus is on the modern period as that subperiod provides the stronger challenge to the asset-pricing models considered. Given the evidence in Tables 7 and 8, we limit our attention to versions of the CAPM, the 2-beta ICAPM, and the 3-beta ICAPM where the zero-beta rate is left unconstrained, though we report results for all ten models in the Appendix.

First, we show that our three-beta model not only describes the cross section of characteristic-sorted portfolios but also can explain the average returns on risk-sorted portfolios. The first three columns of Table 9 price the six risk-sorted portfolios described in Table 6 Panel B in conjunction with six of the 25 size- and book-to-market-sorted portfolios of Table 6 Panel A (the low, medium, and high BE/ME portfolios within the small and large ME quintiles) in the modern period. We find that the three-beta ICAPM is not rejected by the data while the CAPM is rejected. Importantly, the relatively high R^2 for the zero-beta rate version of the volatility ICAPM (68%) is not disproportionately due to characteristic-sorted portfolios, as the R^2 for the risk-sorted subset (80%) is not only comparable to but actually larger than the R^2 for the characteristic-sorted subset (68%).[24] Though the two-beta ICAPM is not rejected, the R^2 is much lower. Appendix Figure 3 provides a graphical summary of these results.[25]

Characteristic and risk-sorted test assets

The next three columns of Table 9 Panel A price the 18 BE/ME-*ivol*-$\widehat{\beta}_{\Delta VAR}$-sorted portfolios described in Table 6 Panel C. These portfolios present an interesting challenge to our model as they incorporate the idiosyncratic risk anomaly of Ang, Hodrick, Xing, and Zhang (2006). We find that the three-beta ICAPM is not rejected by the data while both the CAPM and the two-beta ICAPM are strongly rejected. In particular, the three-beta ICAPM delivers a relatively high R^2 of 49% while the corresponding R^2s of the CAPM and the two-beta ICAPM are less than 15% (with the CAPM picking a negative premium for market beta). As a consequence, our model can also explain the puzzling idiosyncratic risk effect in the cross section of average returns. We provide a visual summary of this success in Appendix Figure 4.

[24] We report the breakdown of R^2 across the risk-sorted and characteristic-sorted subsets in the Internet Appendix.

[25] For completeness, we also examine the ability of our model to explain risk-sorted portfolios in the early period. Internet Appendix Table 4 prices the six risk-sorted portfolios described in Table 5 Panel B in conjunction with six of the 25 size- and book-to-market-sorted portfolios of Table 5 Panel A (the low, medium, and high BE/ME portfolios within the small and large ME quintiles) in the early period. We continue to find that the three-beta ICAPM modestly improves pricing relative to both the Sharpe-Lintner and Black versions of the CAPM. Moreover, the relatively high R^2 (57%) is not disproportionately due to characteristics-sorted portfolios, as the R^2 for the risk-sorted subset (68%) is again not only comparable to but actually larger than the R^2 for the characteristics-sorted subset (51%). Internet Appendix Figure 2 shows this success graphically.

Non-equity test assets

Our three-beta model goes some way towards explaining the average returns on non-equity portfolios designed to be highly correlated with aggregate volatility risk, namely the S&P 100 index straddles of Coval and Shumway (2001). We first calculate the expected return on the straddle portfolio based on the estimate of the zero-beta rate volatility ICAPM in Table 8. The contributions to expected quarterly return from the straddle's cash-flow, discount-rate, and volatility betas are -2.14%, -1.37%, and -3.15% respectively. As the average quarterly realized return on the straddle is -21.66%, an equity-based estimate of the three-beta model explains roughly 31% of the realized straddle premium.

We then price the straddle in conjunction with both equity and fixed income portfolios. Note that if we were simply to add the straddle position to our analysis, its extreme in-sample average return and volatility would dominate the cross-sectional regression. Therefore, we rescale the straddle by forming a portfolio of the straddle and the Treasury Bill. In particular, we hold one seventh of the portfolio in the straddle and the remaining six sevenths in the T-Bill.[26] The last three columns of Table 9 Panel A show that our intertemporal CAPM with stochastic volatility is rejected at the 5-percent level when we price the joint cross-section of equity, bond, and straddle returns. Both the CAPM and the two-beta ICAPM are also strongly rejected. Nevertheless, both the R^2 and the pricing measure improve significantly as one moves from the CAPM to the three-beta ICAPM. Appendix Figure 5 provides a visual summary of the CAPM and ICAPM results, illustrating the difficult challenge posed by the extremely low realized returns on the straddle portfolio.

The first three columns of Table 9 Panel B price the cross section of currency portfolios sorted on interest rates. A large literature studies the properties of the currency carry trade, a strategy that exploits the forward premium puzzle of Fama (1984) by buying high interest-rate currencies and shorting low interest-rate currencies. The two-beta Campbell and Vuolteenaho (2004) model produces a similar R^2 as the CAPM. However, the implied risk aversion estimate is quite large at 14.3.

The proportion of variation explained by the three-beta ICAPM is roughly 82%, with a risk aversion estimate of 6.9 (once again equal to the theoretical maximum). Our intertemporal CAPM with stochastic volatility is not rejected at the 5-percent level by either set of critical values. Appendix Figure 6 provides a visual summary of the relative pricing ability of the CAPM and the volatility ICAPM for currency portfolios.

The second three columns of Table 9 Panel B price the cross section of variance forwards. We find that average returns on one-month variance forwards are negative and statistically significant. At longer horizons, average returns generally increase, and, for some months, become statistically greater than zero. The three-beta model is not rejected by the data; the

[26] We chose these weights as the volatility of the resulting portfolio is roughly similar to the volatility of the other assets included in the test. However, our results are qualitatively unaffected by reasonable variation in these proportions.

explained variation is nearly 84% with a risk aversion estimate of 6.0. Appendix Figure 7 provides a visual summary of these results.

Thus, Table 9 confirms the general success of the three-beta ICAPM in pricing test portfolios beyond the 25 size- and book-to-market-sorted portfolios. Moreover, though the CAPM and the two-beta ICAPM do fit the cross-section in some of the tests of Table 9, a careful look reveals that the pricing parameters for these models are not stable across the various sets of test portfolios, whereas the pricing parameters for the three-beta ICAPM generally are. We demonstrate this fact by also reporting in Table 9 the constrained R^2 – the R^2 that results when a model's parameters are fixed at the corresponding estimates of Table 8. In general, we find that for the CAPM and the two-beta ICAPM, this constrained R^2 is dramatically lower than the unconstrained counterpart. However, for the three-beta ICAPM, there is much less of a difference between the constrained and unconstrained estimates.

7 Historical Implications of the Model

As a further check on the reasonableness of our model, we now look at its implications for the interpretation of US financial history. As a first exercise, we ask what our model implies for the history of investors' marginal utility.

Figure 10 (third panel) plots the time-series of the smoothed combined shock $\gamma N_{CF} - N_{DR} - \frac{1}{2}\omega N_V$ based on the estimate of the zero-beta model for the modern period (Table 8). The correlation of this shock with the associated N_{CF} is 0.82. Similarly, the correlation of this shock with the associated $-N_{DR}$ is 0.21. Finally, the correlation of this shock with the associated N_V is -0.71. Figure 10 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 Campbell, Giglio, and Polk (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.

7.1 Actual versus implied consumption growth

The news terms estimated in our model were obtained using returns data only, after substituting consumption out of the Euler equation. However, the model estimates have implications for the investor's consumption process. In this section we explore the relation between the innovations to consumption growth and cash flows implied by our model and innovations in observed real nondurables and services consumption growth Δc, real dividend growth Δd, and real corporate earnings growth Δe. We focus on innovations because they are relevant for cross-sectional asset pricing, and they can be directly linked to our three news terms.

As we report in the Appendix, our model implies that consumption innovations are given by

$$\Delta c_{t+1} - E_t \Delta c_{t+1} = (r_{t+1} - E_t r_{t+1}) - (\psi - 1) N_{DR,t+1} - (\psi - 1) \frac{1}{2} \frac{\omega}{1 - \gamma} N_{V,t+1}, \tag{30}$$

while cash flow news is determined by equation (17). We can easily compute both series by appropriately combining the estimated VAR innovations with our estimates for γ and ω, respectively 6.9 and 20.9. Note, however, that the EIS parameter ψ, which affects the implied consumption growth innovations, is not pinned down by our VAR estimation. Therefore, we calibrate it to three different values, 0.5, 1.0, and 1.5.

We start by comparing the implied consumption innovations ($\Delta c_{t+1} - E_t \Delta c_{t+1}$) to observed innovations in real log consumption growth, real log dividend growth and real log earnings growth. We construct consumption growth using nondurable and services data from the BEA, and obtain long-term dividend and earnings data from Professor Robert Shiller's website. We construct consumption, dividends and earnings innovations by taking the residuals of an AR(1) regression for each series. To maximize the length of our time series and to avoid the issue of seasonality in dividends and earnings, we work here with yearly data, from 1930 to 2011. All the results are robust to using smoothed quarterly data for the available periods.

Panel A of Table 10 reports the standard deviations of six series: innovations in actual consumption, dividends and earnings growth, and implied innovations in consumption under the three calibration for ψ (0.5, 1, 1.5). The table shows that the implied consumption innovation series are more volatile than actual consumption innovations. The fact that aggregate consumption growth is smooth is well known; as pointed out for example in Malloy, Moskowitz and Vissing-Jørgensen (2009), aggregate consumption growth may not be the right benchmark for our asset pricing model because it includes consumption of non-stockholder consumers.[27] However, the table also shows that implied consumption growth is about as volatile as earnings and dividend growth innovations (depending on ψ).

Panel B of Table 10 reports the correlations between the consumption innovation series implied by our model and the observed consumption, dividend and earnings growth innovations. The left panel reports the correlations of the raw series, while the right panel reports correlations of exponentially-weighted moving averages of each series, which are useful to show the low-frequency comovements. The table shows that our implied consumption series are positively correlated with the realizations of consumption, dividend and earnings growth. The strength of the correlations increases in almost all cases when looking at the smoothed series.

To gauge the low-frequency comovements between these series better, Appendix Figure 10 shows the exponentially-weighted moving average of the series of implied consumption in-

[27] In theory, we could test the model directly using data on stockholder consumption, but the time series available spans 20 years, which is not long enough to be used to estimate the news terms and test our model.

novations versus actual consumption, dividends and earnings growth innovations. Given the quite different standard deviations of these series (as reported above), we have standardized all of them before plotting. We use a value of $\psi = 0.5$ when calibrating implied consumption innovations in this graph, so that the volatility of implied consumption innovation matches that of dividend growth; results are similar for $\psi = 1.0$ and 1.5. The figure shows that the time variation in consumption, dividend and earnings growth innovations aligns well with our implied consumption series, for example capturing most of the major booms and busts.

Finally, we turn to the implied N_{CF} series. Panel C of Table 10 reports correlations between our implied N_{CF} series and future long-run consumption growth (looking at the next 5, 10 and 15 years). The left side of the table uses the raw N_{CF} series, while the right side of the table uses the exponentially smoothed N_{CF} as a predictor. The table shows that the implied N_{CF} is indeed positively correlated with future consumption, dividends and earnings growth in almost all cases, and the relations are generally stronger for longer horizons and when using the smoothed N_{CF} series. Overall, the table suggests that our N_{CF} news term is indeed related to the actual consumption process.

8 Robustness

Table 11 examines the robustness of our findings. Where appropriate, we include in bold font our baseline model as a benchmark. Panel A shows results using various subsets of variables in our baseline VAR. These results indicate that including both DEF and PE are generally essential for our finding of a negative β_V for HML, consistent with the importance of these two variables in long-run volatility forecasting. Moreover, successful zero-beta-rate volatility ICAPM pricing in the modern period requires PE, DEF, and VS in the VAR. The results in Panel A also show that the positive volatility beta of the aggregate stock index in the modern period is due to the inclusion of PE and DEF in the VAR. This finding makes sense once one is convinced (and the long-horizon regressions of Table 4 make a strong case) that, controlling for DEF, high PE forecasts high volatility in the future. Since the market will certainly covary positively (and quite strongly) with the PE shock, one should expect this component of volatility news to be positive and an important determinant of the market's volatility beta.

Panel B presents results based on different estimation methods for the VAR. These methods include OLS, WLS but with OLS betas, two different bounds on the maximum ratio of WLS weights, a single-stage approach where the weights are proportional to $RVAR$ rather than $EVAR$, and a partial VAR where we throw out in each regression those variables with t-statistics under 1.0 (in an iterative fashion, starting with the weakest t-statistic first). These results show that our major findings (a negative β_V for HML and successful zero-beta rate ICAPM pricing in both time periods) are very robust to using different methods.

In Panel C, we vary the way in which we estimate realized variance. In the second, fifth, and sixth columns of the Table, we estimate the VAR using annual data. Thus our estimate of realized variance reflects information over the entire year. In columns three and five, we compute the realized variance of monthly returns rather than the realized variance of daily returns as in our benchmark specification. In the fourth and six columns, we simply sum squared monthly returns. Across Panel C, the R^2s of the zero-beta rate ICAPM remain high in the modern period.

In Panel D, we alter the set of variables included in the VAR as a response to the concern of Chen and Zhao (2009) that VAR-based forecasts are sensitive to this choice. (See also Engsted, Pedersen, and Tanggaard 2012 for a clarifying discussion of this issue.) We first explore different ways to measure the market's valuation ratio. In the second column of the Table, we replace PE with $PE_{\mathrm{Re}al}$ where we construct the price-earnings ratio by deflating both the price and the earnings series by the CPI before taking their ratio. In the third column, we use the log price-dividend ratio, PD, instead of PE. In column four, we replace PE with $PE_{\mathrm{Re}al}$ and the CPI inflation rate, $INFL$. Panel D also explores adding two additional state variables. In column five, we add CAY (Lettau and Ludvigson (2001)) to the VAR as CAY is known to be a strong predictor of future market returns. Column six adds the quarterly $FIGARCH$ forecast to the VAR as Table 4 Panel B documents that GARCH-based methods are useful predictors of future market return variance.

Column seven adds the volatility of the term spread ($TYVol$) to the list of state variables based on the evidence in Fornari and Mele (2011) that this variable contains information about time-varying expected returns. In particular, following Fornari and Mele (2011), we add the mean absolute monthly change in the term spread over the previous twelve months.[28] In results reported in the online appendix, we show that $TYVol$ is not incrementally important for forecasting either the first or second moment of the real market return. In fact, even if we exclude some or all of PE, TY, DEF, or VS from the VAR, $TYVol$ never comes in significantly. However, $TYVol$ does help forecast DEF. Nevertheless, adding $TYVOL$ to the VAR does not qualitatively change the conclusions of the pricing tests. In total, this Panel confirms that our finding of a negative β_V for HML and successful zero-beta rate ICAPM pricing in both time periods is generally robust to these variations.

In Panel E, we study the out-of-sample properties of our model. In particular, we estimate our baseline VAR on an expanding window, using the estimates in each window to generate news terms for the quarter ahead. Since the interesting pricing results are in the modern period, our initial window is the 1926:2 to 1963:2 period, so that the first out-of-sample news realizations occur in 1963:3, corresponding to the first data point in the modern period. We find that the out-of-sample news terms are extremely correlated with their in-sample counterparts. Specifically, the correlations are 0.81, 0.97, and 0.82 for the cash-flow, discount-rate, and volatility news terms respectively. Table 11 Panel E documents that the pricing of the out-of-sample news terms is very consistent with the full-sample results.[29]

[28] Results are robust to measuring the volatility of the monthly term spread over the last year instead.

[29] Internet Appendix Figure 8 provides a visual summary of these out-of-sample pricing results for the CAPM and three-beta ICAPM.

In Panel F, we explore using alternative proxies for the wealth portfolio. In particular, we replace the market returns with the return on a delevered market portfolio that combines Treasury Bills and the market in various constant proportions. By doing so, we are able to assess how varying the volatility of this central series affects our results. The three specific delevered portfolios we examine have 80%, 60%, or 40% invested in the market. We find that the cross-sectional fit of our model remains high, with R^2s in the early period essentially unaffected and R^2s in the modern period increasing in all but the extreme delevered portfolio. Perhaps not surprisingly, the estimated risk aversion parameter increases as the degree of delevering increases. In the modern period, delevering the market portfolio by 20% results in a risk aversion estimate of 9.99, delevering by 40% requires risk aversion of 14.51, and delevering by 80% generates a risk aversion estimate of 18.44.

Panel G reports the results when we vary ρ and the excess zero-beta rate. One might argue that our excess zero-beta rate estimate of 112 basis points a quarter is too high to be consistent with equilibrium. Fortunately, we find that R^2s remain reasonable for excess zero-beta rates that are as low as 50 bps/quarter when ρ takes only a slightly lower value, 0.94.

Panels H and I present information to help us better understand the volatility betas we have estimated for the market as a whole, and for value stocks relative to growth stocks. Panel H reports components of $RMRF$ and HML's β_V in each period (estimated either with WLS or OLS). Specifically, these results use the elements of the vector defined in equation (17) and the corresponding VAR shock to measure how each shock contributes to the β_V in question. Panel H documents, consistent with Panel A, that the excess return on the market has a positive volatility beta in the modern period due in part to the PE state variable. The results in Panel H also show that all of the non-zero components of HML's β_V in the modern period are negative. This finding is comforting as it further confirms that our negative HML beta finding is robust. Panel H also reports OLS estimates of simple betas on $RVAR$ and the 15-year horizon $FIGARCH$ forecast (FIG_{60}) for HML and the excess market return. The HML betas based on these two simple proxies have the same sign as our more sophisticated and more appropriate measure of volatility news. However, conclusions about the relevance of volatility risk for the value effect clearly depend on measuring the long-run component of volatility well.

Finally, Panel I reports time-series regressions of HML on $N_{V,t}$ by itself as well as on all three factors together. We find that $N_{V,t}$ explains over 20% of HML's returns in the modern period. The three news factors together explain slightly over 28%. Thus our model is able to explain not only the cross-sectional variation in average returns of the 25 size- and book-to-market-sorted portfolios of Fama and French (1993) but also a significant amount of time series variation in realized returns on the key factor that they argue is multifactor-minimum-variance (Fama and French, 1996).

9 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 may deteriorate either because expected stock returns decline or because the volatility of stock returns increases. A conservative long-term investor will wish to hedge against both types of changes in investment opportunities; thus, 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 coefficient of relative risk aversion.

Our implementation models the return on the aggregate stock market as one element of a vector autoregressive (VAR) system; the volatility of all shocks to the VAR is another element of the system. The empirical implementation of our VAR 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 sorted both by characteristics (book-market ratios and idiosyncratic volatility) and past volatility betas. We also use the model to fit the average returns on non-equity test assets, including corporate bonds, equity index option straddles, interest-rate sorted currency portfolios, and variance forwards. The explanatory power of the model is quite good across all these sets of test assets, with stable parameter estimates, although the extremely low average return on equity index options remains an empirical challenge.

Our empirical work is limited in two important respects. First, we have assumed that the wealth portfolio of a representative investor can be adequately proxied by a diversified equity portfolio. We have checked that our findings are robust to the use of a delevered equity portfolio, but to the extent that other risky assets, such as corporate bonds or even human capital, are important constituents of wealth, this assumption may be inadequate. It will be of interest to explore stochastic volatility in alternative proxies for the wealth portfolio.

Second, we test only the unconditional implications of the model and do not evaluate its conditional implications. A full conditional test is likely to be a challenging hurdle for the model. To see why, recall that we assume a rational long-term investor always holds 100%

of his or her assets in equities. However, time-variation in real stock returns generally gives the long-term investor an incentive to shift the relative weights on cash and equity, unless real interest rates and market volatility move in exactly the right way to make the equity premium proportional to market volatility. Although we do not explicitly test whether this is the case, early papers by Campbell (1987) and Harvey (1989, 1991) and a recent review by Lettau and Ludvigson (2010) reject this proportionality restriction.

One way to support the assumption of constant 100% equity investment is to invoke binding leverage constraints. Indeed, in the modern sample, the Black (1972) version of our three-beta model is consistent with this interpretation as the estimated difference between the zero-beta and risk-free rates is positive, statistically significant, and economically large. However, while our estimates imply that the leverage constraint is unconditionally binding for a long-horizon investor given the risk characteristics of the market portfolio in the modern sample, they do not imply that it is always conditionally binding in every period.

Both these limitations are opportunities for future research. And our model does directly answer the interesting microeconomic question: Are there reasonable preference parameters that would make a long-term investor, constrained to invest 100% in equity, content to hold the market rather than tilting towards value stocks or other high-return stock portfolios? Our answer is clearly yes.

References

Adrian, Tobias and Joshua Rosenberg, 2008, “Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk”, *Journal of Finance* 63:2997–3030.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys, 2003, “Modeling and Forecasting Realized Volatility”, *Econometrica* 71:579–625.

Andersen, Torben, Tim Bollerslev, Peter F. Christoffersen and Francis X. Diebold, 2006, “Volatility and Correlation Forecasting”, in Graham Elliott, Clive W.J. Granger and Allan Timmermann (eds), *Handbook of Economic Forecasting, Vol I.*

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, “The Cross-Section of Volatility and Expected Returns”, *Journal of Finance* 61:259–299.

Baillie, Richard T., Tim Bollerslev and Hans Ole Mikkelsen, 1996, “Fractionally integrated generalized autoregressive conditional heteroskedasticity”, *Journal of Econometrics* 74:3–30.

Bali, Turan and Nusret Cakici, 2009, “Idiosyncratic Volatility and the Cross Section of Expected Stock Returns”, *Journal of Financial and Quantitative Analysis* 43:29–58.

Bansal, Ravi and Amir Yaron, 2004, “Risks for the Long Run”, *Journal of Finance* 59:1481–1509.

Bansal, Ravi, Dana Kiku and Amir Yaron, 2012, “An Empirical Evaluation of the Long-Run Risks Model for Asset Prices”, *Critical Finance Review* 1:183–221.

Bansal, Ravi, Dana Kiku, Ivan Shaliastovich and Amir Yaron, 2013, “Volatility, the Macroeconomy and Asset Prices”, *Journal of Finance* forthcoming.

Barinov, Alexander, 2013, “Idiosyncratic Volatility, Growth Options, and the Cross-Section of Returns”, working paper, University of Georgia.

Barndorff-Nielsen, Ole E. and Neil Shephard, 2002, “Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models”, *Journal of the Royal Statistical Society B*, 64(2):253–280.

Beeler, Jason and John Y. Campbell, 2012, “The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment”, *Critical Finance Review* 1:141–182.

Black, Fisher, 1972, “Capital Market Equilibrium with Restricted Borrowing”, *Journal of Business* 45:444–454.

Black, Fisher, 1976, “Studies of Stock Price Volatility Changes”, *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistics Section,* Washington 177–181.

Bollerslev, Tim, 1986, "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 31:307–327.

Bollerslev, Tim and Hans Ole Mikkelsen, 1996, "Modeling and Pricing Long Memory in Stock Market Volatility", *Journal of Econometrics* 73:151–184.

Breeden, Douglas T., 1979, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities", *Journal of Financial Economics* 7:265–296.

Calvet, Laurent and Adlai Fisher, 2007, "Multifrequency News and Stock Returns", *Journal of Financial Economics* 86:178–212.

Campbell, John Y., 1987, "Stock Returns and the Term Structure", *Journal of Financial Economics* 18:373–399.

Campbell, John Y., 1993, "Intertemporal Asset Pricing Without Consumption Data", *American Economic Review* 83:487–512.

Campbell, John Y., 1996, "Understanding Risk and Return", *Journal of Political Economy* 104:298–345.

Campbell, John Y., Stefano Giglio, and Christopher Polk, 2013, "Hard Times", *Review of Asset Pricing Studies* 3:95–132.

Campbell, John Y., Stefano Giglio, Christopher Polk, and Robert Turley, 2014, "Appendix to An Intertemporal CAPM with Stochastic Volatility", available online at http://scholar.harvard.edu/campbell/publications.

Campbell, John Y. and Ludger Hentschel, 1992, "No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns", *Journal of Financial Economics* 31:281–318.

Campbell, John Y., Christopher Polk, and Tuomo Vuolteenaho, 2010, "Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns" *Review of Financial Studies* 23:305–344.

Campbell, John Y. and Robert J. Shiller, 1988a, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors", *Review of Financial Studies* 1:195–228.

Campbell, John Y. and Robert J. Shiller, 1988b, "Stock Prices, Earnings, and Expected Dividends", *Journal of Finance* 43:661–676.

Campbell, John Y. and Glen B. Taksler, 2003, "Equity Volatility and Corporate Bond Yields", *Journal of Finance* 58:2321–2349.

Campbell, John Y. and Tuomo Vuolteenaho, 2004, "Bad Beta, Good Beta", *American Economic Review* 94:1249–1275.

Chacko, George and Luis M. Viceira, 2005, “Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets”, *Review of Financial Studies* 18:1369–1402.

Chen, Joseph, 2003, “Intertemporal CAPM and the Cross Section of Stock Returns”, unpublished paper, University of California Davis.

Chen, Zhanhui and Ralitsa Petkova, 2014, Does Idiosyncratic Volatility Proxy for Risk Exposure?, *Review of Financial Studies* 25:2745–2787.

Chen, Long and Xinlei Zhao, 2009, “Return Decomposition”, *Review of Financial Studies* 22:5213–5249.

Christiansen, Charlotte, Maik Schmeling and Andreas Schrimpf, 2012, “A Comprehensive Look at Financial Volatility Prediction by Economic Variables”, *Journal of Applied Econometrics* 27: 956-977.

Christie, Andrew, “The Stochastic Behavior of Common Stock Variances – Value, Leverage, and Interest Rates Effects”, *Journal of Financial Economics* 10:407–432.

Cochrane, John H., 2008, “The Dog That Did Not Bark: A Defense of Return Predictability”, *Review of Financial Studies* 21:1533-1575.

Cochrane, John H., 2011, “Presidential Address: Discount Rates”, *Journal of Finance* 66:1047–1108.

Coval, Josh, and Tyler Shumway, 2001, “Expected Option Returns”, *Journal of Finance* 56:983–1009.

Daniel, Kent and Sheridan Titman, 1997, “Evidence on the Characteristics of Cross-sectional Variation in Common Stock Returns”, *Journal of Finance* 52:1–33.

Daniel, Kent and Sheridan Titman, 2012, “Testing Factor-Model Explanations of Market Anomalies”, *Critical Finance Review* 1:103–139.

Darolles, Serge, Christian Gourieroux, and Joann Jasiak, 2006, “Structural Laplace Transform and Compound Autoregressive Models”, *Journal of Time Series Analysis* 27, 477–504.

Davis, James L., Eugene F. Fama, and Kenneth R. French, “Characteristics, Covariances, and Average Returns: 1929 to 1997”, *Journal of Finance* 55:389–406.

Duffee, Greg, 2005, “Time-Variation in the Covariance Between Stock Returns and Consumption Growth”, *Journal of Finance* 60:1673–172.

Duffie, Darrell and Kenneth J. Singleton, 1997, “An Econometric Model of the Term Structure of Interest Rate Swap Yields”, *Journal of Finance* 52: 1287-1321.

Engle, Robert F, 1982, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50: 987–1007.

Engle, Robert F., Eric Ghysels and Bumjean Sohn, 2013, "Stock Market Volatility and Macroeconomic Fundamentals", *Review of Economics and Statistics* 95: 776–797.

Engsted, Tom, Thomas Q. Pedersen, and Carsten Tanggaard, 2012, "Pitfalls in VAR Based Return Decompositions: A Clarification", *Journal of Banking and Finance* 36: 1255-1265..

Epstein, Lawrence and Stanley Zin, 1989, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica* 57:937–69.

Epstein, Lawrence, Emmanuel Farhi, and Tomasz Strzalecki, 2013, "How Much Would You Pay to Resolve Long-Run Risk?", unpublished paper, Boston University and Harvard University.

Eraker, Bjorn, 2008, "Affine General Equilibrium Models", *Management Science* 54:2068–2080.

Eraker, Bjorn and Ivan Shaliastovich, 2008, "An Equilibrium Guide to Designing Affine Pricing Models", *Mathematical Finance* 18:519–543.

Eraker, Bjorn and Wenyu Wang, 2011, "Dynamic Present Values and the Intertemporal CAPM", University of Wisconsin working paper.

Fama, Eugene F., 1984, "Forward and Spot Exchange Rates", *Journal of Monetary Economics* 14:319–38.

Fama, Eugene F. and Kenneth R. French, 1992, "The Cross-Section of Expected Stock Returns", *Journal of Finance* 47: 427-465.

Fama, Eugene F. and Kenneth R. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics* 33:3–56.

Fama, Eugene F. and Kenneth R. French, 1996, "Multifactor Explanations of Asset Pricing Anomalies", *Journal of Finance* 51:55–84.

Fornari, Fabio and Antonio Mele, 2011, "Financial Volatility and Economic Activity", unpublished paper, University of Lugano.

French, Kenneth, G. William Schwert, and Robert F. Stambaugh, 1987, "Expected Stock Returns and Volatility", *Journal of Financial Economics* 19:3–29.

Fu, Fangjian, 2009, "Idiosyncratic Risk and the Cross-Section of Expected Stock Returns", *Journal of Financial Economics* 91:24–37.

Garcia, Rene, Eric Renault, and A. Semenov, 2006, "Disentangling Risk Aversion and Intertemporal Substitution", *Finance Research Letters* 3:181–193.

Ghysels, Eric, Andrew C. Harvey and Eric Renault, 1996, "Stochastic Volatility", in G.S. Maddala and C.R. Rao eds. *Handbook of Statistics Vol. 14*, 119–191, North-Holland.

Hansen, Lars Peter, 2012, "Dynamic Valuation Decomposition Within Stochastic Economies", *Econometrica* 80:911–967.

Hansen, Lars Peter, 2014, "Uncertainty Inside and Outside Economic Models", unpublished Nobel Lecture, University of Chicago.

Hansen, Lars Peter, John C. Heaton, J. Lee, and Nicholas Roussanov, 2007, "Intertemporal Substitution and Risk Aversion", in J.J. Heckman and E.E. Leamer eds. *Handbook of Econometrics Vol. 6A*, 3967–4056, North-Holland.

Hansen, Lars Peter, John C. Heaton, and Nan Li, 2008, "Consumption Strikes Back? Measuring Long-Run Risk", *Journal of Political Economy* 116:260–302.

Harvey, Campbell, 1989, "Time-Varying Conditional Covariances in Tests of Asset Pricing Models", *Journal of Financial Economics* 24:289–317.

Harvey, Campbell, 1991, "The World Price of Covariance Risk", *Journal of Finance* 46:111–157.

Heston, Steven L., 1993, "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options", *Review of Financial Studies* 6:327–343.

Hull, John, Mirela Predescu and Alan White, 2004, "The Relationship Between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements," *Journal of Banking and Finance* 28:2789-2811.

Kandel, Shmuel and Robert Stambaugh, 1991, "Asset Returns and Intertemporal Preferences", *Journal of Monetary Economics* 27:39–71.

Krishnamurthy, Arvind and Annette Vissing-Jorgensen, 2012, "The Aggregate Demand for Treasury Debt", *Journal of Political Economy* 120:233–267.

Lettau, Martin and Sydney C. Ludvigson, 2010, "Measuring and Modeling Variation in the Risk-Return Trade-off", Chapter 11 in Yacine Ait-Sahalia and Lars Peter Hansen eds. *Handbook of Financial Econometrics*, Elsevier, 617–690.

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2010, "A Skeptical Appraisal of Asset Pricing Tests", *Journal of Financial Economics* 96:175–194.

Liu, Jun, 2007, "Portfolio Selection in Stochastic Environments", *Review of Financial Studies* 20:1–39.

Lustig, Hanno, Nikolai Roussanov, and Adrien Verdelhan, 2011, "Common Risk Factors in Currency Markets", *Review of Financial Studies* 24:3731–3777.

Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-Jørgensen, 2009, "Long-Run Stockholder Consumption Risk and Asset Returns", *Journal of Finance*, 64: 2427–2479.

McQuade, Timothy J., 2012, "Stochastic Volatility and Asset Pricing Puzzles", Harvard University working paper.

Meddahi, Nour and Eric Renault, 2004, "Temporal Aggregation of Volatility Models", *Journal of Econometrics* 119:355–379.

Merton, Robert C., 1973, "An Intertemporal Capital Asset Pricing Model", *Econometrica* 41:867–887.

Paye, Bradley, 2012, "Deja Vol: Predictive Regressions for Aggregate Stock Market Volatility using Macroeconomic Variables", *Journal of Financial Economics* 106: 527–546.

Poon, Ser-Huang, and Clive W. J. Granger, 2003, "Forecasting Volatility in Financial Markets: A Review", *Journal of Economic Literature* 41:478–539.

Restoy, Ferdo, and Philippe Weil, 1998, "Approximate Equilibrium Asset Prices", NBER Working Paper 6611.

Restoy, Ferdo, and Philippe Weil, 2011, "Approximate Equilibrium Asset Prices", *Review of Finance* 15:1–28.

Schwert, William, 1989, "Why Does Stock Market Volatility Change Over Time?", *Journal of Finance* 44:1115–1153.

Sohn, Bumjean, 2010, "Stock Market Volatility and Trading Strategy Based Factors", unpublished paper, Georgetown University.

Tallarini, T., 2000, "Risk Sensitive Real Business Cycles", *Journal of Monetary Economics* 45:507–532.

Tauchen, George and Robert Hussey, 1991, "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models", *Econometrica* 59:317–396.

Table 1: VAR Estimation

The table shows the WLS parameter estimates for a first-order VAR model. The state variables in the VAR include the log real return on the 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 S&P 500's price to the S&P 500's ten-year moving average of earnings (PE), the term yield spread (TY) in percentage points, measured as the difference between the log yield on the ten-year US constant-maturity bond and the log yield on the three-month US Treasury Bill, 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 stages respectively and are then shrunk to equal weights so that the maximum ratio of actual weights used is less than or equal to five. Additionally, the forecasted values for both $RVAR$ and $EVAR$ are constrained to be positive. In Panels A and B, the first seven columns report coefficients on an intercept and the six explanatory variables, and the remaining column 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 of the table reports the correlation ("Corr/std") and autocorrelation ("Autocorr.") matrices of both the unscaled and scaled shocks from the second-stage VAR; the correlation matrix reports shock standard deviations on the diagonal. The sample period for the dependent variables is 1926.3-2011.4, 342 quarterly data points.

Panel A: Forecasting Quarterly Realized Variance ($RVAR_{t+1}$)

Constant	$r_{M,t}$	$RVAR_t$	PE_t	TY_t	DEF_t	VS_t	$R^2\%$
-0.020	-0.004	0.394	0.006	0.000	0.006	0.001	36.88%
(0.008)	(0.005)	(0.064)	(0.002)	(0.001)	(0.001)	(0.002)	

Panel B: VAR Estimates

Second stage	Constant	$r_{M,t}$	$EVAR_t$	PE_t	TY_t	DEF_t	VS_t	$R^2\%$
$r_{M,t+1}$	0.219	0.057	1.249	-0.054	0.004	-0.010	-0.032	2.85%
	(0.119)	(0.068)	(2.276)	(0.034)	(0.008)	(0.023)	(0.035)	
$EVAR_{t+1}$	-0.016	-0.002	0.440	0.005	0.000	0.004	0.002	58.93%
	(0.007)	(0.001)	(0.064)	(0.002)	(0.000)	(0.001)	(0.002)	
PE_{t+1}	0.154	0.138	1.136	0.955	0.004	-0.012	-0.015	94.34%
	(0.114)	(0.066)	(2.178)	(0.032)	(0.007)	(0.022)	(0.033)	
TY_{t+1}	-0.047	-0.097	5.091	0.030	0.820	0.166	0.004	76.56%
	(0.548)	(0.337)	(11.342)	(0.158)	(0.036)	(0.112)	(0.156)	
DEF_{t+1}	0.191	-0.383	6.597	-0.056	0.000	0.834	0.067	88.96%
	(0.264)	(0.151)	(4.856)	(0.073)	(0.017)	(0.050)	(0.075)	
VS_{t+1}	0.138	0.075	3.048	-0.017	-0.004	-0.004	0.939	93.94%
	(0.107)	(0.064)	(2.085)	(0.030)	(0.007)	(0.021)	(0.031)	

Panel C: Correlations and Standard Deviations

Corr/std	r_M	$EVAR$	PE	TY	DEF	VS
			unscaled			
r_M	0.106	-0.488	0.907	-0.022	-0.489	-0.036
$EVAR$	-0.488	0.004	-0.575	-0.074	0.645	0.121
PE	0.907	-0.575	0.099	-0.011	-0.601	-0.064
TY	-0.022	-0.074	-0.011	0.561	0.006	-0.024
DEF	-0.489	0.645	-0.601	0.006	0.290	0.316
VS	-0.036	0.121	-0.064	-0.024	0.316	0.086
			scaled			
r_M	1.137	-0.484	0.904	-0.043	-0.383	0.023
$EVAR$	-0.484	0.045	-0.561	-0.069	0.627	0.088
PE	0.904	-0.561	1.043	-0.033	-0.488	0.004
TY	-0.043	-0.069	-0.033	6.493	0.018	-0.033
DEF	-0.383	0.627	-0.488	0.018	2.727	0.261
VS	0.023	0.088	0.004	-0.033	0.261	0.992

Autocorr.	$r_{M,t+1}$	$EVAR_{t+1}$	PE_{t+1}	TY_{t+1}	DEF_{t+1}	VS_{t+1}
			unscaled			
$r_{M,t}$	-0.074	0.092	-0.067	0.047	0.100	0.045
$EVAR_t$	0.071	-0.153	0.083	-0.126	-0.183	-0.087
PE_t	-0.086	0.177	-0.151	0.070	0.221	0.093
TY_t	-0.046	0.075	-0.029	-0.088	0.081	0.050
DEF_t	0.152	-0.124	0.186	-0.157	-0.311	-0.147
VS_t	0.022	-0.034	0.020	-0.076	-0.080	-0.097
			scaled			
$r_{M,t}$	0.002	0.045	-0.004	0.009	0.007	-0.006
$EVAR_t$	0.060	-0.102	0.073	-0.082	-0.120	-0.060
PE_t	-0.012	0.125	-0.077	0.027	0.109	0.027
TY_t	-0.036	0.067	-0.028	-0.058	0.073	0.039
DEF_t	0.094	-0.083	0.123	-0.111	-0.218	-0.107
VS_t	0.018	-0.031	0.009	-0.044	-0.066	-0.083

Table 2: VAR Specification Test

The table reports the results of regressions forecasting the squared second-stage residuals from the VAR estimated in Table 1 with $EVAR_t$. 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, 342 quarterly data points.

Heteroskedastic Shocks

Squared, second-stage, unscaled residual	Constant	$EVAR_t$	R^2%
$r_{M,t+1}$	-0.003	1.912	19.78%
	[0.004]	[0.309]	
$EVAR_{t+1}$	0.000	0.004	5.86%
	[0.000]	[0.001]	
PE_{t+1}	-0.004	1.937	19.61%
	[0.004]	[0.310]	
TY_{t+1}	0.205	15.082	1.67%
	[0.085]	[7.323]	
DEF_{t+1}	-0.117	27.841	26.12%
	[0.045]	[3.718]	
VS_{t+1}	0.004	0.472	5.47%
	[0.002]	[0.138]	

Table 3: 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 VAR model of Table 1. The upper-left section of the table shows the covariance matrix of the news terms. The upper-right section shows the correlation matrix of the news terms with standard deviations on the diagonal. The lower-left section shows the correlation of shocks to individual state variables with the news terms. The lower-right section 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 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 S&P 500's price to the S&P 500's ten-year moving average of earnings. TY is the term yield spread in percentage points, measured as the difference between the log yield on the ten-year US constant-maturity bond and the log yield on the three-month US Treasury Bill. 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.

News cov.	N_{CF}	N_{DR}	N_V	News corr/std	N_{CF}	N_{DR}	N_V
N_{CF}	0.00213	-0.00042	-0.00026	N_{CF}	0.046	-0.101	-0.221
	(0.00075)	(0.00108)	(0.00028)		(0.007)	(0.228)	(0.257)
N_{DR}	-0.00042	0.00823	-0.00021	N_{DR}	-0.101	0.091	-0.091
	(0.00108)	(0.00261)	(0.00063)		(0.228)	(0.014)	(0.363)
N_V	-0.00026	-0.00021	0.00067	N_V	-0.221	-0.091	0.026
	(0.00028)	(0.00063)	(0.00029)		(0.257)	(0.363)	(0.007)

Shock corr.	N_{CF}	N_{DR}	N_V	Functions	N_{CF}	N_{DR}	N_V
r_M shock	0.523	-0.901	-0.019	r_M shock	0.924	-0.076	-0.013
	(0.210)	(0.036)	(0.335)		(0.030)	(0.030)	(0.016)
$EVAR$ shock	-0.056	0.434	0.452	$RVAR$ shock	-0.368	-0.368	1.289
	(0.151)	(0.113)	(0.159)		(1.068)	(1.068)	(0.553)
PE shock	0.180	-0.967	-0.090	PE shock	-0.856	-0.856	0.189
	(0.239)	(0.037)	(0.357)		(0.165)	(0.165)	(0.087)
TY shock	0.104	0.078	-0.113	TY shock	0.010	0.010	-0.004
	(0.152)	(0.108)	(0.230)		(0.013)	(0.013)	(0.007)
DEF shock	-0.160	0.490	0.741	DEF shock	-0.009	-0.009	0.078
	(0.198)	(0.121)	(0.238)		(0.037)	(0.037)	(0.020)
VS shock	-0.435	-0.179	0.566	VS shock	-0.244	-0.244	0.103
	(0.187)	(0.139)	(0.263)		(0.128)	(0.128)	(0.065)

Table 4: Forecasting Long-Horizon Realized Variance

The table studies the estimates of long-run variance implied by the VAR model of Table 1. Panel A reports the WLS parameter estimates of constrained regressions forecasting the annualized discounted sum of future $RVAR$ over the next 40 quarters ($4 * \sum_{k=1}^{40} \rho^{(k-1)} RVAR_{t+k} / \sum_{k=1}^{40} \rho^{(k-1)}$). The forecasting variables include the VAR state variables defined in Table 1, the corresponding annualized long-horizon forecast implied from estimates of the VAR in Table 1 (VAR_{40}) as well as FIGARCH (FIG_{40}) and two-factor EGARCH (EG_{40}) models estimated from the full sample of daily returns. r_M is the log real return on the 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 S&P 500's price to the S&P 500's ten-year moving average of earnings. TY is the term yield spread in percentage points, measured as the difference between the log yield on the ten-year US constant-maturity bond and the log yield on the three-month US Treasury Bill. 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. PEO is PE orthogonalized to DEF and $DEFO$ is DEF orthogonalized to PE. Initial WLS weights are inversely proportional to the corresponding FIG_{40} long-horizon forecast except in those regressions involving VAR_{40} or EG_{40} forecasts, where the corresponding VAR_{40} or EG_{40} long-horizon forecast is used instead. Newey-West standard errors estimated with lags corresponding to twice the number of overlapping observations are in square brackets. The sample period for the dependent variable is 1930.1-2011.4. Panel B of the table reports summary statistics for realized variance ($RVAR_h$), the corresponding forecasts from the VAR (VAR_h), and the prices of variance swaps (VIX_h^2) at various horizons h. Panel C of the table shows regressions forecasting $LHRVAR_h$ with VAR_h and VIX_h^2. In this Panel, we set ρ to 1 when calculating $LHRVAR_h$ and VAR_h. Newey-West t-statistics that take into account overlapping observations are in brackets.

Panel A: Forecasting 10-year Realized Variance $(4 * \sum_{h=1}^{40} \rho^{(h-1)} RVAR_{t+h} / \sum_{k=1}^{40} \rho^{(h-1)})$

	Constant	r_M	$RVAR$	PE	TY	DEF	VS	VAR_{40}	EG_{40}	FIG_{40}	PEO	$DEFO$	$R^2\%$
1	-0.066	-0.008	0.095	0.024	0.000	0.013	0.001						56.67%
	[0.017]	[0.005]	[0.030]	[0.005]	[0.001]	[0.002]	[0.002]						
2	-0.009							0.989					47.69%
	[0.007]							[0.256]					
3	-0.067								1.458				40.01%
	[0.006]								[0.269]				
4	-0.006									0.987			37.31%
	[0.006]									[0.177]			
5	-0.106	-0.010	0.018	0.023	0.000	0.011	0.001		0.792				58.96%
	[0.021]	[0.005]	[0.022]	[0.004]	[0.001]	[0.002]	[0.002]		[0.250]				
6	-0.075	-0.011	-0.023	0.023	0.000	0.010	-0.001			0.776			60.13%
	[0.016]	[0.005]	[0.021]	[0.005]	[0.001]	[0.001]	[0.001]			[0.222]			
7	-0.016							0.780		0.480			53.92%
	[0.006]							[0.243]		[0.228]			
8	-0.010					0.002		0.943					48.94%
	[0.006]					[0.003]		[0.266]					
9	-0.006			0.009									-0.53%
	(0.026)			(0.009)									
10	0.012					0.008							21.75%
	(0.005)					(0.004)							
11	-0.052										0.025		29.36%
	(0.014)										(0.005)		
12	0.002											0.018	50.60%
	(0.003)											(0.004)	
13	-0.070			0.025		0.017							51.42%
	(0.019)			(0.006)		(0.004)							

Panel B: Comparing VAR_h and VIX_h^2

	$h=1$	$h=2$	$h=3$	$h=4$
mean				
$RVAR$	0.048	0.047	0.047	0.047
VAR_h	0.044	0.044	0.044	0.044
VIX_h^2	0.059	0.058	0.058	0.058
standard deviation				
$RVAR$	0.066	0.057	0.051	0.046
VAR_h	0.021	0.018	0.017	0.015
VIX_h^2	0.042	0.036	0.035	0.034
correlation				
(VAR_h,VIX_h^2)	0.72	0.68	0.66	0.65

Panel C: Forecasting $LHRVAR_h$ with VAR_h and VIX_h^2

$h=1$			$h=2$			$h=3$			$h=4$		
Constant	VAR_h	VIX_h^2	Constant	VAR_h	VIX_h^2	Constant	VAR_h	VIX_h^2	Constant	VAR_h	VIX_h^2
-0.018	1.490		0.002	1.050		0.009	0.853		0.016	0.707	
[-0.94]	[3.79]		[0.12]	[3.28]		[0.67]	[2.84]		[0.98]	[2.42]	
0.009		0.657	0.021		0.457	0.027		0.345	0.032		0.267
[0.62]		[3.30]	[2.31]		[2.09]	[2.97]		[2.18]	[2.79]		[1.97]
-0.016	1.102	0.265	0.002	0.890	0.112	0.010	0.798	0.039	0.016	0.718	-0.008
[-0.86]	[1.94]	[0.95]	[0.18]	[2.99]	[0.53]	[0.70]	[2.34]	[0.23]	[0.97]	[2.08]	[-0.05]

Table 5: Cash-flow, Discount-rate, and Variance Betas in the Early Sample

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 (Panel A) and six risk-sorted portfolios (Panel B). "Growth" denotes the lowest BE/ME, "Value" the highest BE/ME, "Small" the lowest 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 VAR state variables, where the weights are according to λ_V as estimated in Table 3, 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 for the 1931:3-1963:2 period using weighted least squares where the weights are the same as those used to estimate the VAR.

Panel A: 25 ME- and BE/ME-sorted portfolios

$\widehat{\beta}_{CF}$	Growth		2		3		4		Value		Diff	
Small	0.44	[0.13]	0.41	[0.11]	0.39	[0.10]	0.41	[0.10]	0.43	[0.10]	-0.01	[0.06]
2	0.29	[0.07]	0.33	[0.09]	0.33	[0.08]	0.36	[0.08]	0.41	[0.10]	0.12	[0.04]
3	0.29	[0.08]	0.27	[0.08]	0.32	[0.09]	0.32	[0.08]	0.44	[0.12]	0.15	[0.05]
4	0.25	[0.07]	0.26	[0.07]	0.30	[0.08]	0.33	[0.08]	0.43	[0.11]	0.18	[0.05]
Large	0.22	[0.07]	0.22	[0.07]	0.25	[0.08]	0.32	[0.10]	0.38	[0.29]	0.16	[0.04]
Diff	-0.22	[0.07]	-0.19	[0.05]	-0.14	[0.04]	-0.09	[0.03]	-0.05	[0.03]		

$\widehat{\beta}_{DR}$	Growth		2		3		4		Value		Diff	
Small	1.13	[0.15]	1.12	[0.16]	1.09	[0.17]	1.07	[0.16]	1.05	[0.16]	-0.07	[0.07]
2	0.85	[0.11]	0.94	[0.14]	0.91	[0.14]	0.92	[0.16]	1.06	[0.13]	0.21	[0.08]
3	0.86	[0.13]	0.77	[0.09]	0.88	[0.11]	0.86	[0.11]	1.07	[0.15]	0.21	[0.09]
4	0.66	[0.07]	0.75	[0.10]	0.78	[0.09]	0.85	[0.14]	1.11	[0.16]	0.44	[0.13]
Large	0.67	[0.08]	0.63	[0.08]	0.69	[0.11]	0.89	[0.15]	0.89	[0.12]	0.22	[0.13]
Diff	-0.45	[0.14]	-0.49	[0.11]	-0.41	[0.16]	-0.18	[0.13]	-0.17	[0.08]		

$\widehat{\beta}_{V}$	Growth		2		3		4		Value		Diff	
Small	-0.08	[0.05]	-0.09	[0.04]	-0.10	[0.04]	-0.09	[0.04]	-0.10	[0.04]	-0.02	[0.02]
2	-0.05	[0.03]	-0.05	[0.03]	-0.06	[0.03]	-0.07	[0.03]	-0.09	[0.04]	-0.05	[0.02]
3	-0.05	[0.03]	-0.03	[0.02]	-0.05	[0.03]	-0.06	[0.03]	-0.10	[0.04]	-0.05	[0.02]
4	-0.01	[0.02]	-0.03	[0.02]	-0.04	[0.03]	-0.06	[0.04]	-0.10	[0.05]	-0.09	[0.03]
Large	-0.01	[0.02]	-0.02	[0.02]	-0.05	[0.04]	-0.08	[0.04]	-0.07	[0.03]	-0.06	[0.03]
Diff	0.07	[0.04]	0.07	[0.02]	0.04	[0.02]	0.01	[0.02]	0.02	[0.02]		

Panel B: 6 risk-sorted portfolios

$\widehat{\beta}_{CF}$	Lo $\widehat{b}_{r_M}$		2		Hi $\widehat{b}_{r_M}$		Diff	
Lo $\widehat{b}_{VAR}$	0.22	[0.07]	0.33	[0.09]	0.43	[0.12]	0.21	[0.05]
Hi $\widehat{b}_{VAR}$	0.18	[0.06]	0.26	[0.08]	0.36	[0.10]	0.17	[0.05]
Diff	-0.04	[0.02]	-0.07	[0.03]	-0.08	[0.02]		

$\widehat{\beta}_{DR}$	Lo $\widehat{b}_{r_M}$		2		Hi $\widehat{b}_{r_M}$		Diff	
Lo $\widehat{b}_{VAR}$	0.61	[0.07]	0.87	[0.11]	1.09	[0.14]	0.48	[0.09]
Hi $\widehat{b}_{VAR}$	0.55	[0.06]	0.76	[0.09]	0.95	[0.11]	0.40	[0.07]
Diff	-0.07	[0.04]	-0.12	[0.06]	-0.14	[0.05]		

$\widehat{\beta}_{V}$	Lo $\widehat{b}_{r_M}$		2		Hi $\widehat{b}_{r_M}$		Diff	
Lo $\widehat{b}_{VAR}$	-0.02	[0.02]	-0.05	[0.03]	-0.07	[0.04]	-0.05	[0.02]
Hi $\widehat{b}_{VAR}$	-0.01	[0.02]	-0.02	[0.02]	-0.04	[0.03]	-0.03	[0.02]
Diff	0.01	[0.01]	0.03	[0.02]	0.03	[0.01]		

Table 6: Cash-flow, Discount-rate, and Variance Betas in the Modern Sample

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 (Panel A); six risk-sorted portfolios (Panel B); 18 BE/ME, IVol, and $\widehat{\beta}_{\Delta VAR}$-sorted portfolios (Panel C); the S&P 100 index straddle portfolio ($STRADDLE$), the Fama-French factors $RMRF$, SMB, HML, high yield ($HYRET$) and investment grade ($IGRET$) bond portfolios (Panel D); five interest-rate-sorted portfolios of Lustig, Roussanov, and Verdelhan (2011) (Panel E); and 12 VIX Forward positions (delevered twenty-fold) (Panel F). "Growth" denotes the lowest BE/ME, "Value" the highest BE/ME, "Small" the lowest 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 VAR state variables, where the weights are according to λ_V as estimated in Table 3, and on the market-return shock. "Diff." is the difference between the extreme cells. In Panel C, $P1$ is the composite portfolio that is long the equal-weight average of the value portfolios and is short the equal-weight average of the growth portfolios. $P2$ is the composite portfolio that is long the high idiosyncratic portfolio and short the low idiosyncratic portfolio for *either* the growth subset *or* the value subset. $P3$ is the portfolio that is long the equal-weight average of the high $\widehat{\beta}_{\Delta VAR}$ portfolios and is short the equal-weight average of the low $\widehat{\beta}_{\Delta VAR}$ portfolios. Bootstrapped standard errors [in brackets] are conditional on the estimated news series. Estimates are based on quarterly data for the 1963:3-2011:4 period in Panels A, B and C, the 1986:1-2011:4 period in Panel D, the 1984:1-2010:1 period in Panel E, and the 1998:4-2011:4 period in Panel F using weighted least squares where the weights are the same as those used to estimate the VAR.

Panel A: 25 ME- and BE/ME-sorted portfolios

$\widehat{\beta}_{CF}$	Growth		2		3		4		Value		Diff	
Small	0.23	[0.06]	0.24	[0.05]	0.24	[0.04]	0.23	[0.04]	0.26	[0.05]	0.03	[0.03]
2	0.22	[0.05]	0.22	[0.04]	0.24	[0.04]	0.24	[0.04]	0.26	[0.05]	0.04	[0.03]
3	0.20	[0.05]	0.22	[0.04]	0.22	[0.04]	0.23	[0.04]	0.24	[0.04]	0.05	[0.03]
4	0.19	[0.04]	0.21	[0.04]	0.22	[0.04]	0.22	[0.04]	0.24	[0.04]	0.05	[0.03]
Large	0.13	[0.03]	0.17	[0.03]	0.16	[0.03]	0.17	[0.03]	0.19	[0.04]	0.05	[0.03]
Diff	-0.10	[0.04]	-0.07	[0.03]	-0.08	[0.02]	-0.06	[0.02]	-0.07	[0.03]		

$\widehat{\beta}_{DR}$	Growth		2		3		4		Value		Diff	
Small	1.31	[0.10]	1.06	[0.08]	0.89	[0.07]	0.83	[0.07]	0.87	[0.09]	-0.44	[0.08]
2	1.21	[0.09]	0.97	[0.07]	0.85	[0.06]	0.76	[0.07]	0.80	[0.08]	-0.42	[0.08]
3	1.14	[0.07]	0.89	[0.05]	0.77	[0.06]	0.72	[0.06]	0.72	[0.07]	-0.42	[0.08]
4	1.03	[0.06]	0.85	[0.05]	0.74	[0.06]	0.72	[0.06]	0.75	[0.07]	-0.28	[0.08]
Large	0.84	[0.05]	0.71	[0.04]	0.60	[0.05]	0.59	[0.06]	0.64	[0.06]	-0.20	[0.06]
Diff	-0.46	[0.10]	-0.35	[0.08]	-0.29	[0.06]	-0.24	[0.07]	-0.23	[0.08]		

$\widehat{\beta}_{V}$	Growth		2		3		4		Value		Diff	
Small	0.18	[0.07]	0.12	[0.06]	0.08	[0.06]	0.07	[0.05]	0.03	[0.07]	-0.15	[0.03]
2	0.19	[0.07]	0.12	[0.06]	0.08	[0.05]	0.06	[0.06]	0.04	[0.06]	-0.15	[0.03]
3	0.19	[0.06]	0.11	[0.05]	0.08	[0.05]	0.04	[0.06]	0.06	[0.04]	-0.13	[0.03]
4	0.17	[0.06]	0.11	[0.05]	0.06	[0.06]	0.05	[0.06]	0.04	[0.06]	-0.13	[0.03]
Large	0.13	[0.05]	0.10	[0.04]	0.06	[0.04]	0.04	[0.05]	0.04	[0.05]	-0.09	[0.02]
Diff	-0.05	[0.03]	-0.02	[0.03]	-0.03	[0.02]	-0.03	[0.02]	0.01	[0.03]		

Panel B: 6 risk-sorted portfolios

$\widehat{\beta}_{CF}$	Lo $\widehat{b}_{r_M}$		2		Hi $\widehat{b}_{r_M}$		Diff	
Lo $\widehat{b}_{VAR}$	0.16	[0.03]	0.17	[0.03]	0.25	[0.05]	0.08	[0.04]
Hi $\widehat{b}_{VAR}$	0.15	[0.02]	0.17	[0.04]	0.23	[0.05]	0.08	[0.04]
Diff	-0.01	[0.02]	0.00	[0.02]	-0.01	[0.02]		

$\widehat{\beta}_{DR}$	Lo $\widehat{b}_{r_M}$		2		Hi $\widehat{b}_{r_M}$		Diff	
Lo $\widehat{b}_{VAR}$	0.55	[0.05]	0.71	[0.06]	1.11	[0.09]	0.56	[0.08]
Hi $\widehat{b}_{VAR}$	0.73	[0.06]	0.95	[0.05]	1.27	[0.09]	0.54	[0.11]
Diff	0.18	[0.07]	0.24	[0.07]	0.16	[0.06]		

$\widehat{\beta}_{V}$	Lo $\widehat{b}_{r_M}$		2		Hi $\widehat{b}_{r_M}$		Diff	
Lo $\widehat{b}_{VAR}$	0.06	[0.05]	0.08	[0.05]	0.12	[0.07]	0.07	[0.03]
Hi $\widehat{b}_{VAR}$	0.11	[0.04]	0.16	[0.04]	0.18	[0.07]	0.07	[0.04]
Diff	0.05	[0.02]	0.08	[0.02]	0.06	[0.02]		

Panel C: 18 BE/ME, IVol, and $\widehat{\beta}_{\Delta VAR}$-sorted portfolios

$\widehat{\beta}_{CF}$	Growth						2						Value					
	Low IVol		2		High IVol		Low IVol		2		High IVol		Low IVol		2		High IVol	
Low $\widehat{\beta}_{\Delta VAR}$	0.15	[0.03]	0.16	[0.04]	0.23	[0.05]	0.18	[0.03]	0.23	[0.04]	0.28	[0.05]	0.20	[0.04]	0.28	[0.05]	0.27	[0.05]
High $\widehat{\beta}_{\Delta VAR}$	0.14	[0.03]	0.17	[0.05]	0.20	[0.06]	0.17	[0.03]	0.21	[0.04]	0.23	[0.05]	0.16	[0.03]	0.24	[0.04]	0.30	[0.06]
	$P1$: 0.08 [0.02]						$P2$: Growth 0.07 [0.03], Value 0.11 [0.03]						$P3$: -0.02 [0.02]					

$\widehat{\beta}_{DR}$	Growth						2						Value					
	Low IVol		2		High IVol		Low IVol		2		High IVol		Low IVol		2		High IVol	
Low $\widehat{\beta}_{\Delta VAR}$	0.67	[0.05]	0.96	[0.08]	1.21	[0.11]	0.54	[0.05]	0.73	[0.06]	1.00	[0.10]	0.53	[0.06]	0.76	[0.08]	0.86	[0.09]
High $\widehat{\beta}_{\Delta VAR}$	0.86	[0.06]	1.17	[0.08]	1.43	[0.14]	0.64	[0.04]	0.96	[0.07]	1.16	[0.10]	0.63	[0.06]	0.92	[0.07]	1.11	[0.10]
	$P1$: -0.21 [0.06]						$P2$: Growth 0.52 [0.11], Value 0.41 [0.08]						$P3$: 0.18 [0.04]					

$\widehat{\beta}_{V}$	Growth						2						Value					
	Low IVol		2		High IVol		Low IVol		2		High IVol		Low IVol		2		High IVol	
Low $\widehat{\beta}_{\Delta VAR}$	0.08	[0.05]	0.11	[0.08]	0.15	[0.09]	0.03	[0.05]	0.05	[0.05]	0.05	[0.09]	0.01	[0.05]	0.02	[0.08]	0.02	[0.07]
High $\widehat{\beta}_{\Delta VAR}$	0.16	[0.04]	0.18	[0.07]	0.20	[0.09]	0.08	[0.03]	0.11	[0.06]	0.17	[0.06]	0.08	[0.04]	0.07	[0.07]	0.10	[0.07]
	$P1$: -0.10 [0.02]						$P2$: Growth 0.06 [0.05], Value 0.02 [0.03]						$P3$: 0.07 [0.01]					

Panel D: Option, equity, and bond portfolios

	$STRADDLE$		$RMRF$		SMB		HML		$HYRET$		$IGRET$	
$\widehat{\beta}_{CF}$	-0.39	[0.27]	0.18	[0.05]	0.04	[0.02]	0.05	[0.03]	0.06	[0.02]	0.00	[0.01]
$\widehat{\beta}_{DR}$	-1.71	[0.47]	0.81	[0.06]	0.19	[0.05]	-0.25	[0.09]	0.26	[0.07]	0.03	[0.03]
$\widehat{\beta}_{V}$	0.38	[0.21]	-0.00	[0.07]	-0.00	[0.02]	-0.12	[0.03]	-0.05	[0.05]	0.01	[0.01]

Panel E: Currency portfolios

	Low r*		2		3		4		High r*		Diff	
$\widehat{\beta}_{CF}$	0.00	[0.02]	-0.03	[0.02]	0.00	[0.02]	0.02	[0.02]	0.02	[0.02]	0.02	[0.02]
$\widehat{\beta}_{DR}$	-0.16	[0.08]	-0.08	[0.10]	-0.09	[0.07]	0.01	[0.08]	-0.01	[0.08]	0.15	[0.09]
$\widehat{\beta}_{V}$	0.01	[0.03]	-0.04	[0.04]	-0.03	[0.02]	-0.02	[0.03]	-0.05	[0.04]	-0.07	[0.06]

Panel F: VIX^2 Forward Positions

Maturity	$\widehat{\beta}_{CF}$		$\widehat{\beta}_{DR}$		$\widehat{\beta}_{V}$	
1	-0.11	[0.04]	-0.33	[0.19]	0.21	[0.14]
2	-0.06	[0.03]	-0.21	[0.09]	0.11	[0.08]
3	-0.03	[0.02]	-0.13	[0.06]	0.07	[0.05]
4	-0.03	[0.02]	-0.11	[0.06]	0.06	[0.04]
5	-0.02	[0.01]	-0.14	[0.11]	0.10	[0.09]
6	-0.03	[0.01]	-0.04	[0.03]	0.04	[0.01]
7	-0.01	[0.01]	-0.06	[0.04]	0.04	[0.02]
8	-0.03	[0.01]	-0.06	[0.04]	0.04	[0.02]
9	-0.04	[0.02]	-0.02	[0.04]	0.04	[0.02]
10	-0.03	[0.03]	-0.02	[0.04]	0.04	[0.02]
11	-0.01	[0.01]	0.00	[0.04]	0.04	[0.02]
12	0.00	[0.02]	0.00	[0.04]	0.03	[0.02]
12-1	0.12	[0.05]	0.38	[0.22]	-0.22	[0.17]

Table 7: Asset Pricing Tests for the Early Sample

The table shows the premia estimated from the 1931:3-1963:2 sample for the CAPM, the 2-beta ICAPM, the 3-beta volatility ICAPM, a factor model where only the $\widehat{\beta}_{DR}$ premium is restricted, and an unrestricted factor model. The test assets are the 25 ME- and BE/ME-sorted portfolios. The first column per model constrains the zero-beta rate (R_{zb}) to equal the risk-free rate (R_{rf}) while the second column allows R_{zb} to be a free parameter. Estimates are from a cross-sectional regression of average simple excess test-asset returns (quarterly in fractions) on an intercept and estimated cash-flow ($\widehat{\beta}_{CF}$), discount-rate ($\widehat{\beta}_{DR}$), and variance betas ($\widehat{\beta}_{V}$). Standard errors and critical values [A] are conditional on the estimated news series and (B) incorporate full estimation uncertainty of the news terms. The test rejects if the pricing error is higher than the listed 5 percent critical value.

Parameter	CAPM		2-beta ICAPM		3-beta ICAPM		Constrained		Unrestricted	
R_{zb} less R_f (g_0)	0	-0.002	0	0.001	0	0.002	0	0.015	0	0.023
% per annum	0%	-0.90%	0%	0.21%	0%	0.86%	0%	6.08%	0%	8.98%
Std. err. A	0	[0.016]	0	[0.014]	0	[0.011]	0	[0.018]	0	[0.022]
Std. err. B	0	(0.016)	0	(0.016)	0	(0.014)	0	(0.017)	0	(0.020)
$\widehat{\beta}_{CF}$ premium (g_1)	0.038	0.040	0.096	0.094	0.086	0.081	0.081	0.009	0.075	0.057
% per annum	15.11%	15.82%	38.33%	37.74%	34.30%	32.52%	32.48%	3.67%	29.88%	22.69%
Std. err. A	[0.015]	[0.024]	[0.054]	[0.079]	[0.038]	[0.048]	[0.062]	[0.106]	[0.125]	[0.129]
Std. err. B	(0.015)	(0.024)	(0.142)	(0.111)	(0.072)	(0.082)	(0.101)	(0.118)	(0.129)	(0.139)
$\widehat{\beta}_{DR}$ premium (g_2)	0.038	0.040	0.016	0.016	0.016	0.016	0.016	0.016	0.018	-0.012
% per annum	15.11%	15.82%	6.40%	6.40%	6.40%	6.40%	6.40%	6.40%	7.28%	-4.78%
Std. err. A	[0.015]	[0.024]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.050]	[0.064]
Std. err. B	(0.015)	(0.024)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.075)	(0.089)
$\widehat{\beta}_{VAR}$ premium (g_3)					-0.053	-0.045	-0.077	-0.227	-0.080	-0.258
% per annum					-21.27%	-17.97%	-30.67%	-90.62%	-31.96%	-103.04%
Std. err. A					[0.048]	[0.061]	[0.186]	[0.237]	[0.210]	[0.304]
Std. err. B					(0.146)	(0.146)	(0.549)	(0.558)	(0.684)	(0.782)
$\widehat{R^2}$	53.01%	53.12%	54.72%	54.73%	55.65%	55.83%	55.74%	57.94%	55.75%	58.87%
Pricing error	0.024	0.023	0.022	0.022	0.020	0.021	0.020	0.022	0.020	0.023
5% critic. val. A	[0.062]	[0.031]	[0.058]	[0.038]	[0.065]	[0.044]	[0.043]	[0.038]	[0.037]	[0.038]
5% critic. val. B	(0.062)	(0.031)	(0.097)	(0.046)	(0.179)	(0.052)	(0.049)	(0.040)	(0.037)	(0.042)
Implied γ	2.4	2.5	6.0	5.9	5.4	5.1	N/A	N/A	N/A	N/A
Implied ω	N/A	N/A	N/A	N/A	6.7	5.6	N/A	N/A	N/A	N/A

Table 8: Asset Pricing Tests for the Modern Sample

The table shows the premia estimated from the 1963:3-2011:4 sample for the CAPM, the 2-beta ICAPM, the 3-beta volatility ICAPM, a factor model where only the $\widehat{\beta}_{DR}$ premium is restricted, and an unrestricted factor model. The test assets are the 25 ME- and BE/ME-sorted portfolios. The first column per model constrains the zero-beta rate (R_{zb}) to equal the risk-free rate (R_{rf}) while the second column allows R_{zb} to be a free parameter. Estimates are from a cross-sectional regression of average simple excess test-asset returns (quarterly in fractions) on an intercept and estimated cash-flow ($\widehat{\beta}_{CF}$), discount-rate ($\widehat{\beta}_{DR}$), and variance betas ($\widehat{\beta}_{V}$). Standard errors and critical values [A] are conditional on the estimated news series and (B) incorporate full estimation uncertainty of the news terms. The test rejects if the pricing error is higher than the listed 5 percent critical value.

Parameter	CAPM		2-beta ICAPM		3-beta ICAPM		Constrained		Unrestricted	
R_{zb} less R_f (g_0)	0	0.027	0	-0.019	0	0.011	0	-0.004	0	-0.005
% per annum	0%	10.62%	0%	-7.71%	0%	4.50%	0%	-1.66%	0%	-2.00%
Std. err. A	0	[0.014]	0	[0.013]	0	[0.012]	0	[0.013]	0	[0.013]
Std. err. B	0	(0.014)	0	(0.019)	0	(0.015)	0	(0.015)	0	(0.015)
$\widehat{\beta}_{CF}$ premium (g_1)	0.020	-0.004	0.074	0.161	0.047	0.054	0.112	0.128	0.175	0.199
% per annum	7.98%	-1.67%	29.41%	64.39%	18.78%	21.49%	44.65%	51.35%	70.18%	79.55%
Std. err. A	[0.010]	[0.019]	[0.047]	[0.070]	[0.024]	[0.013]	[0.050]	[0.071]	[0.070]	[0.084]
Std. err. B	(0.010)	(0.019)	(0.087)	(0.113)	(0.040)	(0.053)	(0.114)	(0.116)	(0.124)	(0.126)
$\widehat{\beta}_{DR}$ premium (g_2)	0.020	-0.004	0.008	0.008	0.008	0.008	0.008	0.008	-0.018	-0.020
% per annum	7.98%	-1.67%	3.11%	3.11%	3.11%	3.11%	3.11%	3.11%	-7.30%	-7.83%
Std. err. A	[0.010]	[0.019]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.023]	[0.025]
Std. err. B	(0.010)	(0.019)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.054)	(0.055)
$\widehat{\beta}_{VAR}$ premium (g_3)					-0.039	-0.081	-0.094	-0.089	-0.002	0.009
% per annum					-15.51%	-32.47%	-37.65%	-35.60%	-0.72%	3.62%
Std. err. A					[0.039]	[0.024]	[0.063]	[0.069]	[0.092]	[0.094]
Std. err. B					(0.091)	(0.151)	(0.356)	(0.349)	(0.399)	(0.387)
$\widehat{R^2}$	-36.51%	5.22%	25.10%	39.97%	-108.63%	62.74%	73.90%	74.45%	76.46%	77.25%
Pricing error	0.110	0.107	0.058	0.042	0.210	0.042	0.027	0.025	0.026	0.023
5% critic. val. A	[0.050]	[0.034]	[0.061]	[0.056]	[0.503]	[0.101]	[0.051]	[0.037]	[0.046]	[0.031]
5% critic. val. B	(0.049)	(0.033)	(0.096)	(0.083)	(0.492)	(0.119)	(0.104)	(0.078)	(0.065)	(0.049)
Implied γ	2.6	-0.5	9.5	20.7	6.0	6.9	N/A	N/A	N/A	N/A
Implied ω	N/A	N/A	N/A	N/A	10.0	20.9	N/A	N/A	N/A	N/A

Table 9: Asset Pricing Tests for the Modern Sample: Various Sets of Test Assets

The table shows the premia estimated from the 1963:3-2011:4 sample for the CAPM, the 2-beta ICAPM, and the 3-beta volatility ICAPM. Test assets vary across the table and in Panel A are as follows: "Risk-sorted" are six ME- and BE/ME-sorted portfolios and six risk-sorted portfolios, "Idiosyncratic Volatility" are 18 BE/ME, IVol, and $\widehat{\beta}_{\Delta VAR}$-sorted portfolios, and "Straddle, Equity, and Bonds" are the three equity factors of Fama and French (1993), the returns on high yield and investment grade bond portfolios, and the S&P 100 index straddle return from Coval and Shumway (2001). In Panel B, the test assets are "Carry Trade", the five interest-rate-sorted currency portfolios from developed countries of Lustig, Roussanov, and Verdelhan (2011), and "Variance Forward", twelve variance forward positions of maturities from one month to one year. Estimates are from a cross-sectional regression of average simple excess test-asset returns (quarterly in fractions) on an intercept and estimated cash-flow ($\widehat{\beta}_{CF}$), discount-rate ($\widehat{\beta}_{DR}$), and variance betas ($\widehat{\beta}_{V}$). The constrained R^2 row reports the R^2 when model parameters, including R_{zb}, are constrained to be the same as in Table 8. Standard errors and critical values [A] are conditional on the estimated news series and (B) incorporate full estimation uncertainty of the news terms. The test rejects if the pricing error is higher than the listed 5 percent critical value.

Panel A	Risk-sorted			Idiosycratic Volatility			Straddle, Equity, and Bonds		
Parameter	CAPM	2-beta ICAPM	3-beta ICAPM	CAPM	2-beta ICAPM	3-beta ICAPM	CAPM	2-beta ICAPM	3-beta ICAPM
R_{zb} less R_f (g_0)	0.017	-0.004	0.009	0.032	-0.013	0.007	-0.002	-0.004	-0.002
% per annum	6.60%	-1.74%	3.50%	12.83%	-5.34%	2.79%	-0.61%	-1.77%	-0.85%
Std. err. A	[0.009]	[0.011]	[0.011]	[0.012]	[0.013]	[0.012]	[0.002]	[0.003]	[0.003]
Std. err. B	(0.009)	(0.013)	(0.011)	(0.012)	(0.019)	(0.013)	(0.002)	(0.003)	(0.003)
$\widehat{\beta}_{CF}$ premium (g_1)	0.001	0.078	0.054	-0.013	0.114	0.054	0.025	0.144	0.055
% per annum	0.44%	31.12%	21.49%	-5.13%	45.69%	21.49%	9.85%	57.46%	22.11%
Std. err. A	[0.014]	[0.071]	[0.022]	[0.017]	[0.059]	[0.015]	[0.010]	[0.070]	[0.012]
Std. err. B	(0.014)	(0.102)	(0.050)	(0.016)	(0.120)	(0.051)	(0.010)	(0.082)	(0.034)
$\widehat{\beta}_{DR}$ premium (g_2)	0.001	0.008	0.008	-0.013	0.008	0.008	0.025	0.008	0.008
% per annum	0.44%	3.11%	3.11%	-5.13%	3.11%	3.11%	9.85%	3.20%	3.20%
Std. err. A	[0.014]	[0.002]	[0.002]	[0.017]	[0.002]	[0.002]	[0.010]	[0.001]	[0.001]
Std. err. B	(0.014)	(0.002)	(0.002)	(0.016)	(0.002)	(0.002)	(0.010)	(0.001)	(0.001)
$\widehat{\beta}_{VAR}$ premium (g_3)			-0.081			-0.081			-0.084
% per annum			-32.47%			-32.47%			-33.39%
Std. err. A			[0.034]			[0.028]			[0.023]
Std. err. B			(0.131)			(0.145)			(0.106)
$\widehat{R^2}$	8.49%	17.89%	68.23%	14.05%	11.33%	48.62%	50.68%	65.87%	71.24%
constrained $\widehat{R^2}$	-29.21%	-6.62%	57.17%	-2.05%	-4.38%	35.02%	-159.21%	5.84%	16.52%
Pricing error	0.041	0.032	0.016	0.118	0.118	0.060	0.349	0.344	0.292
5% critic. val. A	[0.022]	[0.029]	[0.074]	[0.040]	[0.058]	[0.105]	[0.101]	[0.138]	[0.178]
5% critic. val. B	(0.022)	(0.028)	(0.077)	(0.040)	(0.059)	(0.121)	(0.100)	(0.194)	(0.181)
Implied γ	0.1	10.0	6.9	-1.7	14.7	6.9	3.1	18.0	6.9
Implied ω	N/A	N/A	20.9	N/A	N/A	20.9	N/A	N/A	20.9

Panel B	Carry trade			Variance Forwards		
		2-beta	3-beta		2-beta	3-beta
Parameter	CAPM	ICAPM	ICAPM	CAPM	ICAPM	ICAPM
R_{zb} less R_f (g_0)	0.007	0.005	0.003	0.010	0.010	0.012
% per annum	2.79%	2.17%	1.32%	3.95%	4.06%	4.60%
Std. err. A	[0.005]	[0.005]	[0.004]	[0.003]	[0.004]	[0.004]
Std. err. B	(0.005)	(0.005)	(0.005)	(0.003)	(0.004)	(0.004)
$\widehat{\beta}_{CF}$ premium (g_1)	0.030	0.107	0.052	0.071	0.256	0.070
% per annum	12.17%	42.85%	20.72%	28.34%	102.29%	28.10%
Std. err. A	[0.011]	[0.104]	[0.020]	[0.042]	[0.135]	[0.020]
Std. err. B	(0.011)	(0.083)	(0.031)	(0.042)	(0.147)	(0.051)
$\widehat{\beta}_{DR}$ premium (g_2)	0.030	0.008	0.008	0.071	0.010	0.010
% per annum	12.17%	3.00%	3.00%	28.34%	4.07%	4.07%
Std. err. A	[0.011]	[0.001]	[0.001]	[0.042]	[0.002]	[0.002]
Std. err. B	(0.011)	(0.001)	(0.001)	(0.042)	(0.002)	(0.002)
$\widehat{\beta}_{VAR}$ premium (g_3)	0.0000	0.0000	-0.078	0.0000	0.0000	-0.106
% per annum	0.00%	0.00%	-31.31%	0.00%	0.00%	-42.45%
Std. err. A	[0.0000]	[0.0000]	[0.029]	[0.0000]	[0.0000]	[0.037]
Std. err. B	(0.0000)	(0.0000)	(0.064)	(0.0000)	(0.0000)	(0.172)
$\widehat{R^2}$	59.41%	59.67%	81.59%	83.49%	75.91%	83.61%
constrained $\widehat{R^2}$	-4190.89%	-5329.39%	-476.65%	-738.66%	-768.78%	83.53%
Pricing error	0.008	0.008	0.004	0.209	0.341	0.170
5% critic. val. A	[0.030]	[0.042]	[0.079]	[0.679]	[0.757]	[1.226]
5% critic. val. B	(0.030)	(0.044)	(0.076)	(0.668)	(0.779)	(1.055)
Implied γ	4.1	14.3	6.9	7.0	25.2	6.9
Implied ω	N/A	N/A	20.9	N/A	N/A	20.9

Table 10: Actual and Implied Consumption

The table reports a comparison of the consumption innovations and cash flow news series implied by the model and their data counterparts (obtained using observed consumption, dividends and earnings). We use yearly data between 1930 and 2011. Real consumption growth is constructed from BEA data, real dividend and earnings growth series are obtain from Professor Shiller's website. Innovations in these variables are constructed by taking a residual of an AR(1) regression for each series. Implied cash flow news series are constructed by estimating a yearly version of our baseline VAR. Panel A reports standard deviations of implied consumption innovations (for different values of ψ) and actual consumption, dividends, and earnings innovations. Panel B reports the correlations between implied consumption and actual innovations in consumption, dividends and earnings. The left part of the table reports correlations with the innovations of the raw series. The right side of the table reports correlations with smoothed consumption innovations (using the exponential moving average as in Figure 3, with the same smoothing parameter of weight of 0.08 per quarter, or 0.29 per year). Panel C reports the correlations of the raw (left side) and exponentially-smoothed (right side) N_{CF} series with future cumulative consumption growth, looking 5, 10 and 15 years ahead.

Panel A: Standard Deviations

Δc (actual)	0.019
Δd (actual)	0.108
Δe (actual)	0.291
Δc (implied, $\psi = 0.5$)	0.103
Δc (implied, $\psi = 1$)	0.177
Δc (implied, $\psi = 1.5$)	0.287

Panel B: Implied vs. actual consumption innovations

	Raw series			Smoothed series		
	Δc (actual)	Δd (actual)	Δe (actual)	Δc (actual)	Δd (actual)	Δe (actual)
Δc (implied, $\psi = 0.5$)	0.23	0.19	0.21	0.22	0.39	0.23
Δc (implied, $\psi = 1$)	0.29	0.24	0.23	0.30	0.40	0.26
Δc (implied, $\psi = 1.5$)	0.27	0.23	0.20	0.31	0.36	0.24

Panel C: N_{CF} vs. actual long-run consumption

	Raw N_{CF} series			Smoothed N_{CF} series		
	5-year	10-year	15-year	5-year	10-year	15-year
Long-run Δc	-0.03	-0.06	0.07	-0.07	0.01	0.25
Long-run Δd	0.00	0.05	0.01	0.18	0.18	0.07
Long-run Δe	-0.07	0.04	0.06	0.06	0.22	0.17

Table 11: Robustness

The table provides a variety of robustness tests. When appropriate, the baseline model appears in bold font. Panel A reports the results when only a subset of state variables from the baseline VAR ($D \equiv DEF$, $T \equiv TERM$, $V \equiv VS$, $P \equiv PE$) are used to forecast returns and realized variance. Panel B reports the results when different estimation techniques are used. Panel C reports results as we change the estimate of realized variance. Panel D reports the results when other state variables either replace or are added to the VAR. These variables include the log real PE ratio ($PE_{\mathrm{Re}\,al}$), the log price-dividend ratio (PD), log inflation ($INFL$), CAY, the quarterly $FIGARCH$ variance forecast (FIG), and the term spread volatility ($TYVol$). Panel E reports the modern-period results when out-of-sample versions of the model's news terms are used in the pricing tests. Panel F reports results using delevered market portfolios. Panel G reports results when the excess zero-beta rate is varied from 40 to 86 basis points per quarter. Panel H reports the components of $RMRF$ and HML's $\widehat{\beta}_V$ by re-estimating $\widehat{\beta}_V$ using each component of $\mathbf{e2}'\lambda_V$. Panel I also reports simple loadings of $RMRF$ and HML on $RVAR$ and the 15-year $FIGARCH$ variance forecast. Panel J reports time-series regressions explaining HML with the three news terms described in Table 3.

Panel A: Results Using Various Subsets of the Baseline VAR (r_M and $RVAR$ always included)

	None	D	D/T/V	**ALL**	P/D/V	P/D	P
$\widehat{\gamma}^{Max}$	4.5	3.2	3.1	**6.9**	6.7	8.9	14.9
				Early Period			
				$\widehat{\beta}_V$			
$RMRF$	-0.03	-0.23	-0.20	**-0.03**	-0.03	-0.02	0.08
SMB	-0.01	-0.08	-0.07	**-0.02**	-0.02	-0.02	0.03
HML	0.00	-0.12	-0.14	**-0.06**	-0.06	-0.04	0.06
				Risk-free Rate ICAPM			
$\widehat{\gamma}$	2.1	2.3	2.4	**5.4**	5.4	5.5	8.1
$\widehat{\omega}$	2.0	2.5	3.2	**6.7**	6.9	5.3	12.3
$\widehat{R^2}$	51.35%	51.63%	53.78%	**55.65%**	55.63%	52.14%	42.27%
				Zero-beta Rate ICAPM			
$\widehat{R}_{zb}$ less R_f	-0.08%	-0.20%	-0.10%	**0.21%**	0.22%	-0.10%	0.78%
$\widehat{\gamma}$	2.2	2.3	2.5	**5.1**	5.1	5.7	5.3
$\widehat{\omega}$	2.1	2.8	3.4	**5.6**	5.8	5.7	4.4
$\widehat{R^2}$	51.36%	51.72%	53.81%	**55.83%**	55.82%	52.17%	44.61%
				Modern Period			
				$\widehat{\beta}_V$			
$RMRF$	-0.11	-0.15	-0.07	**0.10**	0.11	0.07	0.00
SMB	-0.03	-0.05	-0.03	**0.02**	0.03	0.01	0.00
HML	0.00	-0.01	-0.10	**-0.11**	-0.11	-0.05	-0.02
				Risk-free Rate ICAPM			
$\widehat{\gamma}$	2.3	2.3	2.6	**6.0**	5.6	8.6	6.15
$\widehat{\omega}$	3.0	2.9	4.7	**10.0**	8.3	21.6	6.34
$\widehat{R^2}$	-37.26%	-37.15%	1.71%	**-108.63%**	-136.46%	-27.78%	-43.34%
				Zero-beta Rate ICAPM			
$\widehat{R}_{zb}$ less R_f	2.12%	2.09%	-0.70%	**1.12%**	1.22%	0.45%	1.60%
$\widehat{\gamma}$	0.0	0.0	3.1	**6.9**	6.7	8.9	0.0
$\widehat{\omega}$	1.3	1.1	12.6	**20.9**	19.4	28.3	0.2
$\widehat{R^2}$	5.04%	6.72%	21.89%	**62.74%**	57.42%	3.38%	-0.48%

Panel B: Results Using Different Estimation Methods

	All OLS	OLS Betas	WLS 3	**WLS 5**	WLS 8	RVAR Weighted	Partial VAR
$\widehat{\gamma}^{Max}$	1.5	6.9	7.0	**6.9**	6.7	5.8	4.1
				Early Period			
				$\widehat{\beta}_V$			
RMRF	-0.08	-0.06	-0.05	**-0.03**	-0.01	-0.02	0.00
SMB	-0.03	-0.03	-0.02	**-0.02**	-0.02	-0.02	-0.02
HML	-0.08	-0.07	-0.07	**-0.06**	-0.05	-0.05	-0.06
				Risk-free Rate ICAPM			
$\widehat{\gamma}$	1.5	4.6	5.1	**5.4**	5.7	5.4	4.1
$\widehat{\omega}$	4.7	4.2	5.7	**6.7**	8.4	8.4	10.6
$\widehat{R^2}$	-178.32%	55.23%	54.95%	**55.65%**	56.50%	56.47%	19.70%
				Zero-beta Rate ICAPM			
$\widehat{R}_{zb}$ less R_f	1.85%	0.29%	0.27%	**0.21%**	0.17%	0.16%	0.80%
$\widehat{\gamma}$	1.5	4.3	4.8	**5.1**	5.5	5.2	4.1
$\widehat{\omega}$	4.7	3.3	4.7	**5.6**	7.3	7.2	9.6
$\widehat{R^2}$	47.74%	55.51%	55.22%	**55.83%**	56.64%	56.59%	58.27%
				Modern Period			
				$\widehat{\beta}_V$			
RMRF	0.07	0.07	0.09	**0.10**	0.12	0.09	0.17
SMB	0.02	0.02	0.02	**0.02**	0.02	0.02	0.04
HML	-0.09	-0.09	-0.10	**-0.11**	-0.11	-0.13	-0.13
				Risk-free Rate ICAPM			
$\widehat{\gamma}$	1.4	6.5	6.3	**6.0**	5.6	5.6	2.4
$\widehat{\omega}$	1.4	13.2	12.1	**10.0**	7.9	9.8	0.9
$\widehat{R^2}$	-356.05%	-76.45%	-81.60%	**-108.63%**	-144.04%	-95.09%	-337.38%
				Zero-beta Rate ICAPM			
$\widehat{R}_{zb}$ less R_f	1.41%	0.91%	0.99%	**1.12%**	1.27%	0.91%	1.88%
$\widehat{\gamma}$	1.5	6.9	7.0	**6.9**	6.7	5.8	4.1
$\widehat{\omega}$	4.7	20.9	21.8	**20.9**	19.3	13.7	10.6
$\widehat{R^2}$	12.56%	58.22%	62.57%	**62.74%**	61.71%	55.28%	51.83%

Panel C: Results Using Different Measures of Realized Variance

	Quarterly Var Daily	Annual Var Daily	Quarterly Var Monthly	Quarterly Sum Monthly	Annual Var Monthly	Annual Sum Monthly
$\widehat{\gamma}^{Max}$	**6.9**	5.8	5.1	5.1	6.2	5.5
			Early Period			
			$\widehat{\beta}_V$			
$RMRF$	**-0.03**	0.21	-0.18	-0.24	0.12	0.19
SMB	**-0.02**	0.07	-0.08	-0.07	0.04	0.05
HML	**-0.06**	0.06	-0.14	-0.15	0.00	0.10
			Risk-free Rate ICAPM			
$\widehat{\gamma}$	**5.4**	3.7	4.0	4.2	5.5	4.6
$\widehat{\omega}$	**6.7**	2.0	3.5	3.1	6.9	3.4
$\widehat{R^2}$	**55.65%**	-991.53%	56.95%	53.85%	-781.40%	-887.87%
			Zero-beta Rate ICAPM			
$\widehat{R}_{zb}$ less R_f	**0.21%**	3.68%	0.11%	0.19%	3.31%	3.51%
$\widehat{\gamma}$	**5.1**	1.5	3.9	4.1	3.7	2.7
$\widehat{\omega}$	**5.6**	0.1	3.3	2.7	1.8	0.5
$\widehat{R^2}$	**55.83%**	16.70%	56.99%	53.96%	18.74%	17.89%
			Modern Period			
			$\widehat{\beta}_V$			
$RMRF$	**0.10**	0.12	0.03	-0.01	0.09	0.10
SMB	**0.02**	-0.02	-0.01	-0.03	-0.03	-0.06
HML	**-0.11**	-0.06	-0.12	-0.14	-0.05	-0.06
			Risk-free Rate ICAPM			
$\widehat{\gamma}$	**6.0**	4.5	5.1	5.1	5.6	5.3
$\widehat{\omega}$	**10.0**	3.9	11.6	8.7	7.1	6.1
$\widehat{R^2}$	**-108.63%**	-361.48%	-27.69%	2.44%	-235.53%	-229.17%
			Zero-beta Rate ICAPM			
$\widehat{R}_{zb}$ less R_f	**1.12%**	1.83%	0.66%	0.49%	1.45%	1.35%
$\widehat{\gamma}$	**6.9**	5.8	5.1	5.1	6.2	5.5
$\widehat{\omega}$	**20.9**	13.4	11.6	8.7	14.0	9.0
$\widehat{R^2}$	**62.74%**	62.14%	56.63%	49.08%	58.73%	54.42%

Panel D: Results Replacing/Adding Other State Variables to the VAR

	PE	PE_{Real}	PD	$INFL$	CAY	FIG	$TYVOL$
$\widehat{\gamma}^{Max}$	**6.9**	8.8	4.5	9.0	14.7	5.4	
Early Period							
$\widehat{\beta}_V$							
$RMRF$	**-0.03**	0.01	-0.12	0.02	0.06	-0.03	-0.01
SMB	**-0.02**	-0.01	-0.05	-0.01	0.01	-0.02	-0.01
HML	**-0.06**	-0.04	-0.10	-0.03	0.00	-0.03	-0.04
Risk-free Rate ICAPM							
$\widehat{\gamma}$	**5.4**	5.4	3.1	5.8	12.4	5.0	5.8
$\widehat{\omega}$	**6.7**	4.3	2.9	5.1	26.8	8.0	8.4
$\widehat{R^2}$	**55.65%**	55.89%	57.77%	55.98%	-1228.81%	56.57%	55.42%
Zero-beta Rate ICAPM							
$\widehat{R}_{zb}$ less R_f	**0.21%**	-0.17%	-0.15%	-0.23%	2.19%	0.10%	0.22%
$\widehat{\gamma}$	**5.1**	5.7	3.2	6.2	14.7	4.9	5.4
$\widehat{\omega}$	**5.6**	4.9	3.2	6.0	56.6	7.3	7.0
$\widehat{R^2}$	**55.83%**	55.97%	57.83%	56.13%	23.45%	56.60%	55.62%
Modern Period							
$\widehat{\beta}_V$							
$RMRF$	**0.10**	0.12	-0.01	0.13	0.06	0.09	0.10
SMB	**0.02**	0.03	-0.01	0.03	0.02	0.03	0.02
HML	**-0.11**	-0.08	-0.09	-0.08	-0.05	-0.07	-0.11
Risk-free Rate ICAPM							
$\widehat{\gamma}$	**6.0**	8.0	4.1	8.0	14.5	4.6	6.1
$\widehat{\omega}$	**10.0**	13.4	8.3	12.8	48.3	5.7	10.3
$\widehat{R^2}$	**-108.63%**	-32.02%	2.59%	-38.01%	17.06%	-183.16%	-99.86%
Zero-beta Rate ICAPM							
$\widehat{R}_{zb}$ less R_f	**1.12%**	0.72%	-0.36%	0.81%	0.35%	1.21%	1.10%
$\widehat{\gamma}$	**6.9**	8.8	4.5	9.0	14.7	5.4	6.9
$\widehat{\omega}$	**20.9**	22.8	15.9	23.4	56.6	13.5	20.8
$\widehat{R^2}$	**62.74%**	26.08%	20.12%	31.92%	35.16%	41.44%	65.68%

Panel E: Modern Period Pricing		
	Full Sample	Out of Sample
	$\widehat{\beta}_V$	
$RMRF$	**0.10**	-0.04
SMB	**0.02**	-0.02
HML	**-0.11**	-0.10
Risk-free Rate ICAPM		
$\widehat{\gamma}$	**6.0**	6.8
$\widehat{\omega}$	**10.0**	20.7
$\widehat{R^2}$	**-108.63%**	52.09%
Zero-beta Rate ICAPM		
$\widehat{R}_{zb}$ less R_f	**1.12%**	0.22%
$\widehat{\gamma}$	**6.9**	6.9
$\widehat{\omega}$	**20.9**	20.1
$\widehat{R^2}$	**62.74%**	59.92%

Panel F: Results Using delevered market portfolios

Equity %	**100%**	80%	60%	40%
$\widehat{\gamma}^{Max}$	**6.9**	10.0	15.1	21.5
		Early Period		
		$\widehat{\beta}_V$		
$RMRF$	**-0.03**	-0.03	-0.03	-0.02
SMB	**-0.02**	-0.02	-0.02	-0.02
HML	**-0.06**	-0.06	-0.05	-0.05
		Risk-free Rate ICAPM		
$\widehat{\gamma}$	**5.36**	7.96	12.24	18.84
$\widehat{\omega}$	**6.65**	17.27	58.58	395.71
$\widehat{R^2}$	**55.65%**	55.45%	54.82%	53.25%
		Zero-beta Rate ICAPM		
$\widehat{R}_{zb}$ less R_f	**0.21%**	0.31%	0.44%	0.71%
$\widehat{\gamma}$	**5.08**	7.41	11.19	16.73
$\widehat{\omega}$	**5.62**	14.09	46.33	291.98
$\widehat{R^2}$	**55.83%**	55.86%	55.69%	55.88%
		Modern Period		
		$\widehat{\beta}_V$		
$RMRF$	**0.10**	0.10	0.10	0.08
SMB	**0.02**	0.02	0.02	0.02
HML	**-0.11**	-0.11	-0.11	-0.09
		Risk-free Rate ICAPM		
$\widehat{\gamma}$	**6.04**	8.69	12.95	17.56
$\widehat{\omega}$	**9.97**	22.70	68.78	329.06
$\widehat{R^2}$	**-108.63%**	-120.66%	-114.82%	-180.16%
		Zero-beta Rate ICAPM		
$\widehat{R}_{zb}$ less R_f	**1.12%**	1.19%	1.04%	1.11%
$\widehat{\gamma}$	**6.91**	9.99	14.51	18.44
$\widehat{\omega}$	**20.88**	45.20	101.91	373.59
$\widehat{R^2}$	**62.74%**	64.59%	63.55%	55.91%

Panel G: Varying ρ and the Excess Zero-beta Rate in the Modern Period

$\rho = 0.94$						
$\widehat{R}_{zb}$ less R_f	0.40%	0.50%	0.60%	0.70%	0.80%	0.91%
$\widehat{\gamma}$	6.7	6.8	6.9	7.0	7.0	7.1
$\widehat{\omega}$	15.3	16.5	18.0	19.9	22.1	22.7
$\widehat{R^2}$	32.51%	44.22%	52.83%	58.84%	62.78%	64.21%
Pricing error	0.062	0.051	0.044	0.039	0.038	0.038

$\rho = 0.95$						
$\widehat{R}_{zb}$ less R_f	0.40%	0.50%	0.60%	0.70%	0.80%	**1.12%**
$\widehat{\gamma}$	6.3	6.3	6.5	6.6	6.7	**6.9**
$\widehat{\omega}$	11.4	12.1	13.0	14.3	15.8	**20.9**
$\widehat{R^2}$	0.48%	18.27%	32.50%	43.46%	51.54%	**62.74%**
Pricing error	0.094	0.076	0.062	0.052	0.046	**0.042**

$\rho = 0.96$						
$\widehat{R}_{zb}$ less R_f	0.40%	0.50%	0.60%	0.70%	0.80%	1.35%
$\widehat{\gamma}$	5.7	5.8	5.9	6.0	6.1	6.7
$\widehat{\omega}$	8.1	8.5	9.0	9.8	10.8	19.1
$\widehat{R^2}$	-42.66%	-18.68%	1.56%	18.21%	31.48%	60.67%
Pricing error	0.138	0.113	0.092	0.076	0.063	0.046

Panel H: Components of and Proxies for $\widehat{\beta}_V$

Early Period				
	$RMRF$		HML	
	WLS	OLS	WLS	OLS
$\widehat{\beta}_V$	-0.03	-0.06	-0.06	-0.07
$\widehat{\beta}_{\lambda_V^1 r_M\ Shock}$	0.00	0.00	0.00	0.00
$\widehat{\beta}_{\lambda_V^2 EVAR\ Shock}$	-0.06	-0.05	-0.02	-0.03
$\widehat{\beta}_{\lambda_V^3 PE\ Shock}$	0.14	0.14	0.06	0.08
$\widehat{\beta}_{\lambda_V^4 TY\ Shock}$	0.00	0.00	0.00	0.00
$\widehat{\beta}_{\lambda_V^5 DEF\ Shock}$	-0.13	-0.13	-0.09	-0.09
$\widehat{\beta}_{\lambda_V^6 VS\ Shock}$	-0.02	-0.02	-0.02	-0.03
$\widehat{\beta}_{RVAR}$		-0.01		1.50
$\widehat{\beta}_{FIGARCH}$		0.02		0.04

Modern Period				
	$RMRF$		HML	
	WLS	OLS	WLS	OLS
$\widehat{\beta}_V$	0.10	0.07	-0.11	-0.09
$\widehat{\beta}_{\lambda_V^1 r_M\ Shock}$	0.00	0.00	0.00	0.00
$\widehat{\beta}_{\lambda_V^2 EVAR\ Shock}$	-0.08	-0.07	-0.01	-0.01
$\widehat{\beta}_{\lambda_V^3 PE\ Shock}$	0.14	0.14	-0.02	-0.02
$\widehat{\beta}_{\lambda_V^4 TY\ Shock}$	0.00	0.00	0.00	0.00
$\widehat{\beta}_{\lambda_V^5 DEF\ Shock}$	-0.02	-0.02	-0.02	-0.01
$\widehat{\beta}_{\lambda_V^6 VS\ Shock}$	0.03	0.03	-0.05	-0.05
$\widehat{\beta}_{RVAR}$		-3.31		-0.51
$\widehat{\beta}_{FIGARCH}$		-0.08		-0.01

Panel I: Time-series Regressions explaining HML

Early Period				
	(1)		(2)	
Intercept	0.01	1.30	0.01	1.99
N_{CF}			0.39	3.78
$-N_{DR}$			0.40	6.69
N_V	-2.00	-6.61	-1.31	-5.02
$\widehat{R^2}$	25.14%		50.91%	

Modern Period				
	(1)		(2)	
Intercept	0.01	2.30	0.01	2.54
N_{CF}			0.25	2.27
$-N_{DR}$			-0.24	-4.69
N_V	-0.99	-7.09	-0.66	-4.46
$\widehat{R^2}$	20.33%		28.45%	

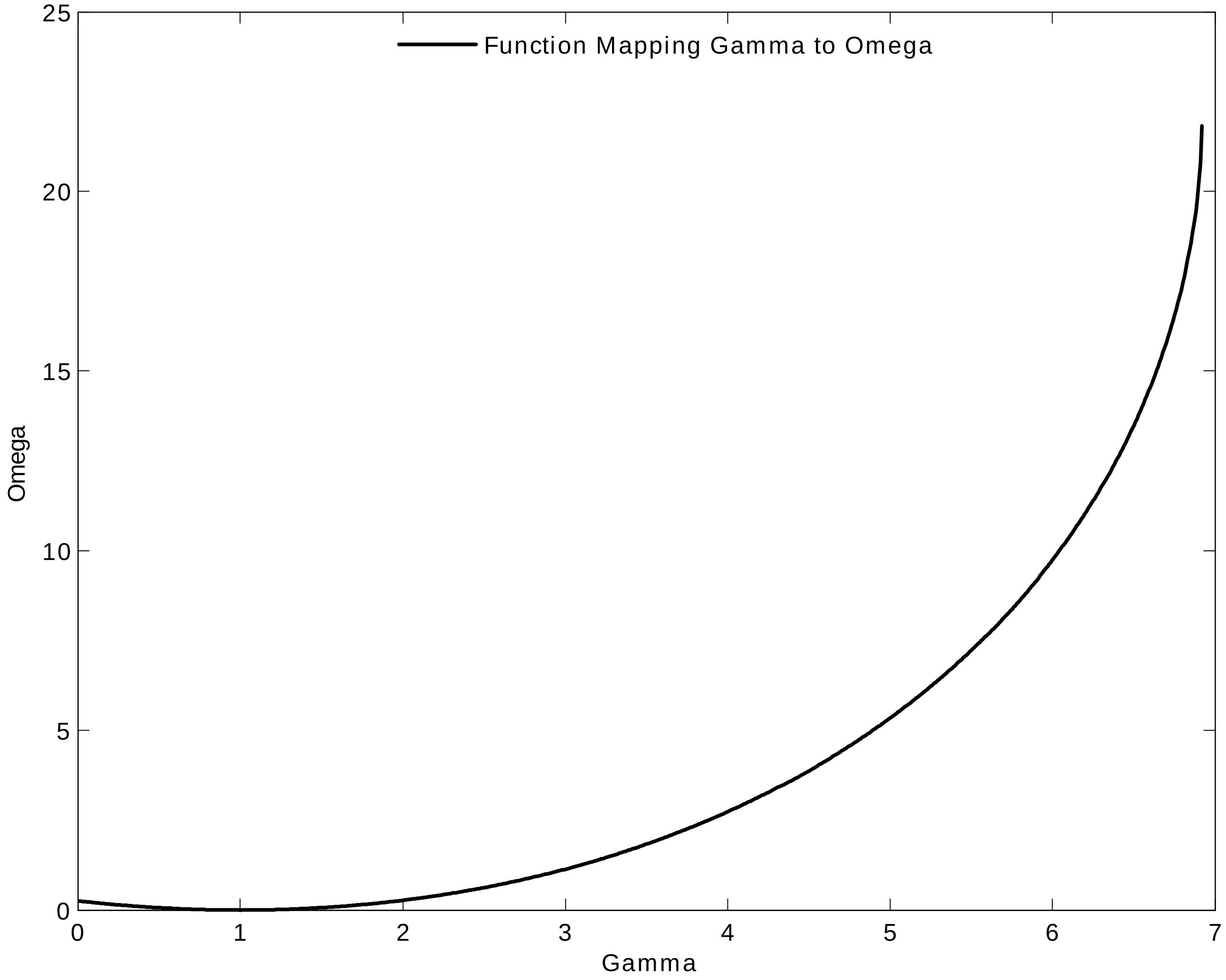

Figure 1: This figure graphs the relation between the parameter γ and the parameter ω described by equation (24). These functions depend on the loglinearization parameter ρ, set to 0.95 per year and the empirically estimated VAR parameters of Table 1. γ is the investor's risk aversion while ω is the sensitivity of news about risk, N_{RISK}, to news about market variance, N_V.

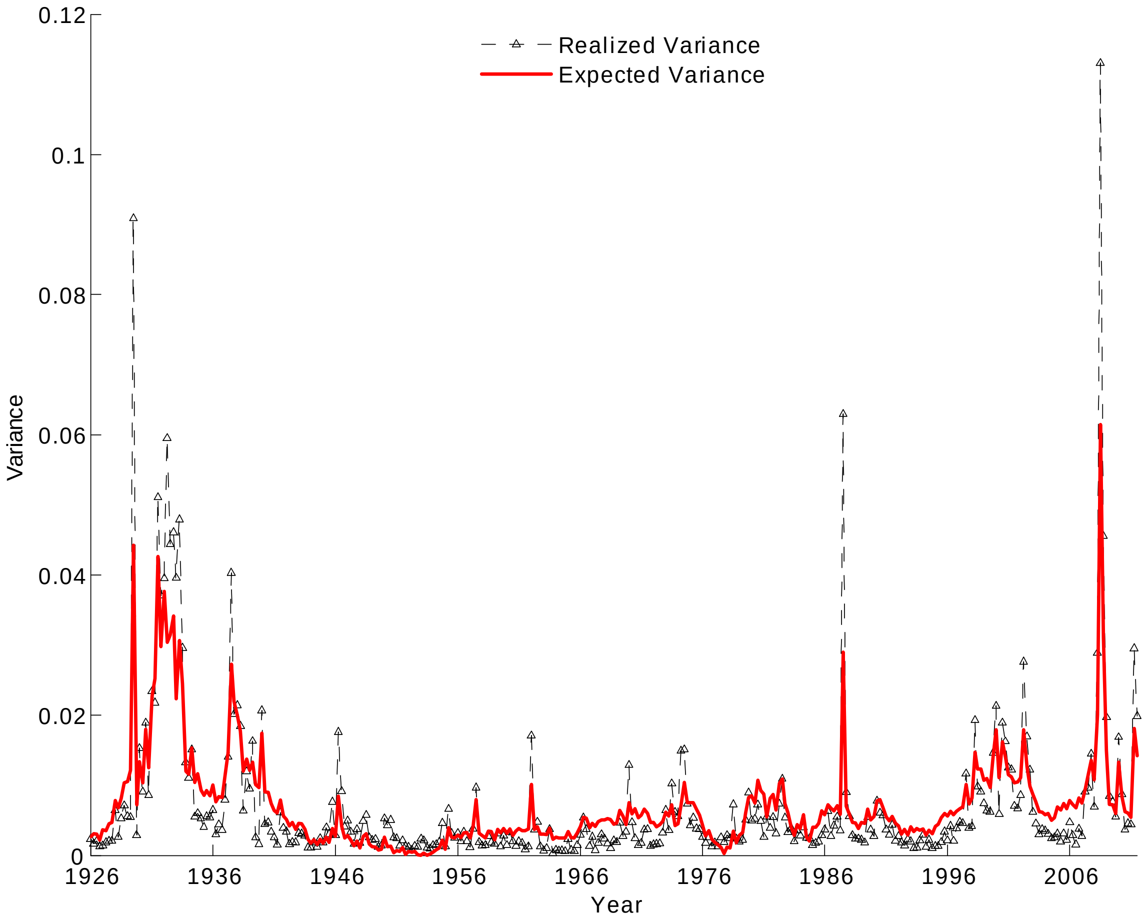

Figure 2: This figure plots quarterly observations of realized within-quarter daily return variance over the sample period 1926:2-2011:4 and the expected variance implied by the model estimated in Table 1 Panel A.

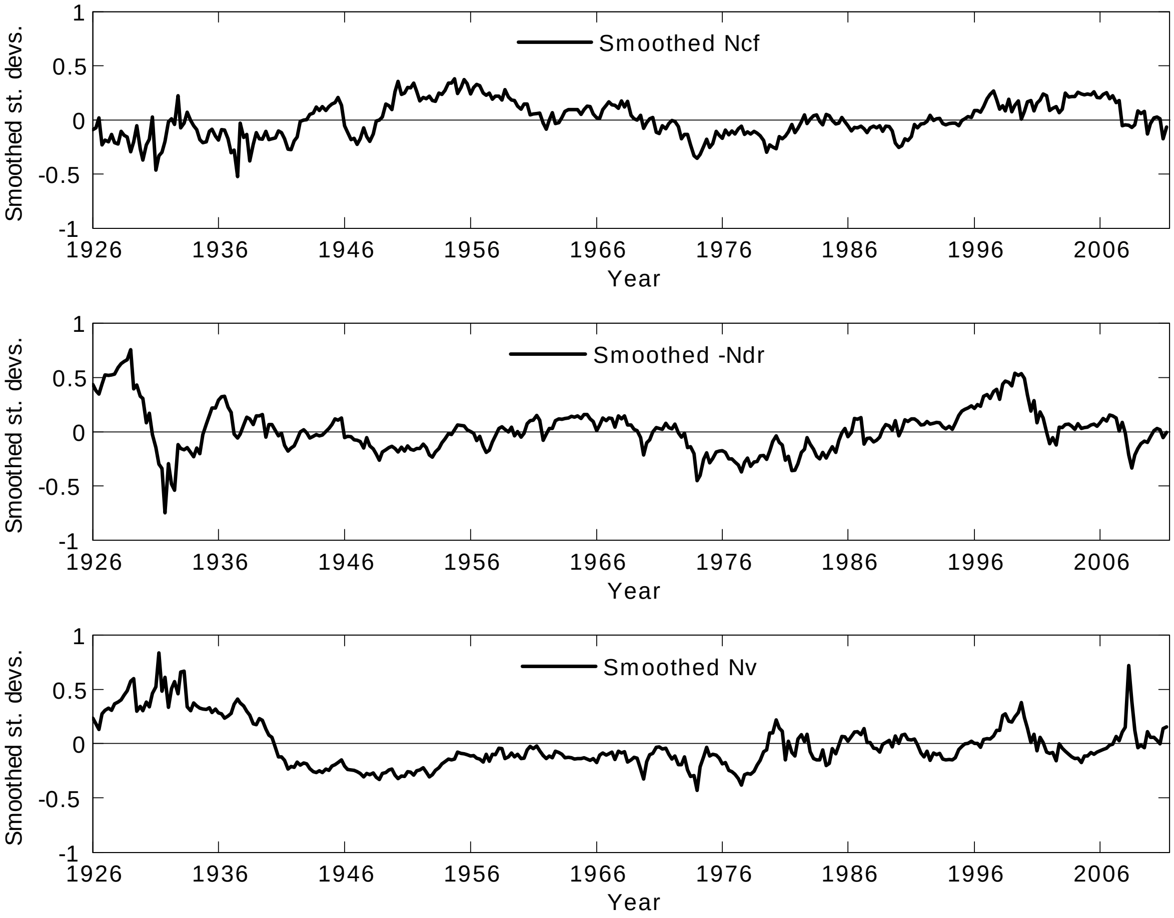

Figure 3: This figure plots normalized cash-flow news, the negative of normalized discount-rate news, and normalized variance news. The series are smoothed with a trailing exponentially-weighted moving average where the decay parameter is set to 0.08 per quarter, and the smoothed news series is generated as $MA_t(N) = 0.08N_t + (1-0.08)MA_{t-1}(N)$. This decay parameter implies a half-life of six years. The sample period is 1926:2-2011:4.

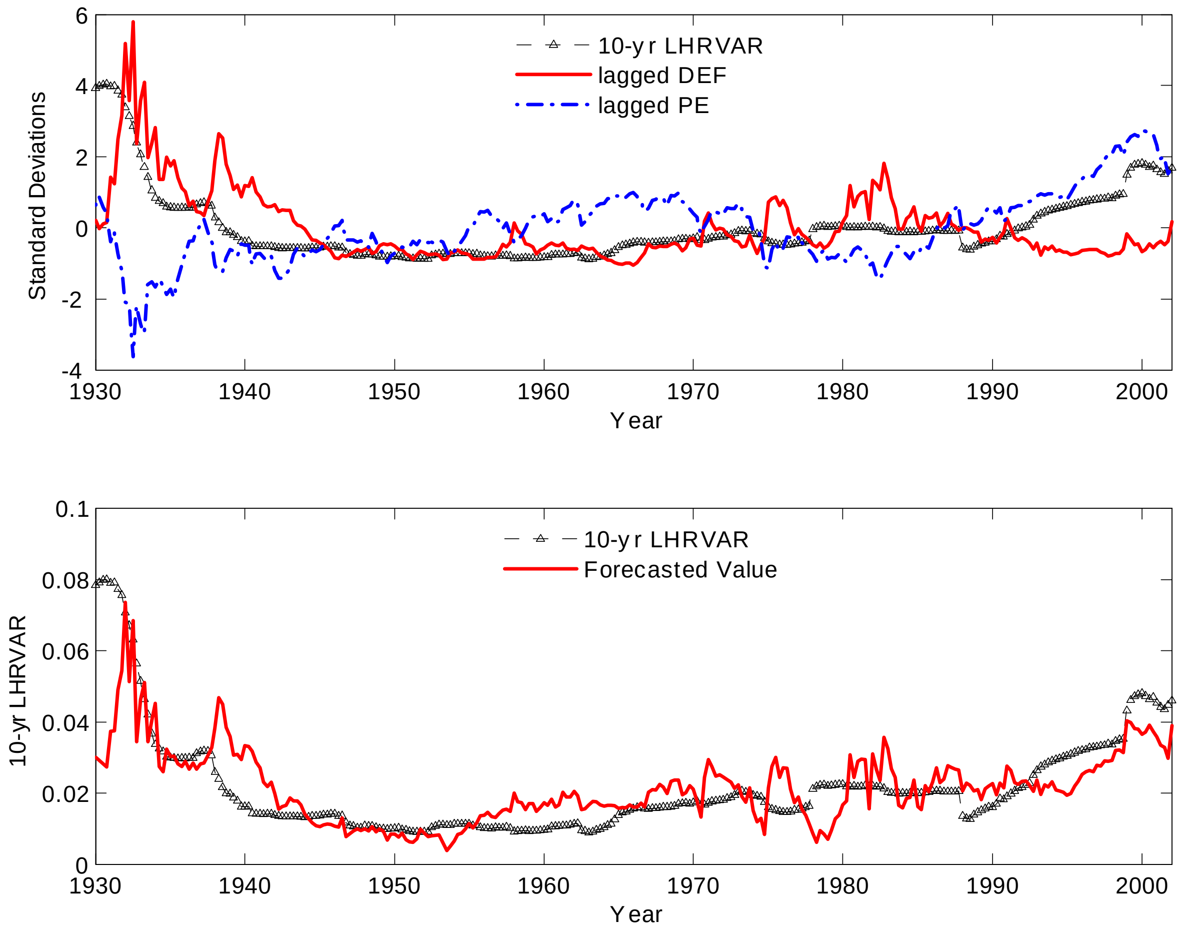

Figure 4: We measure long-horizon realized variance ($LHRVAR$) as the annualized discounted sum of within-quarter daily return variance, $LHRVAR_h = \frac{4*\Sigma_{j=1}^{h}\rho^{j-1}RVAR_{t+j}}{\Sigma_{j=1}^{h}\rho^{j-1}}$. Each panel of this figure plots quarterly observations of ten-year realized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

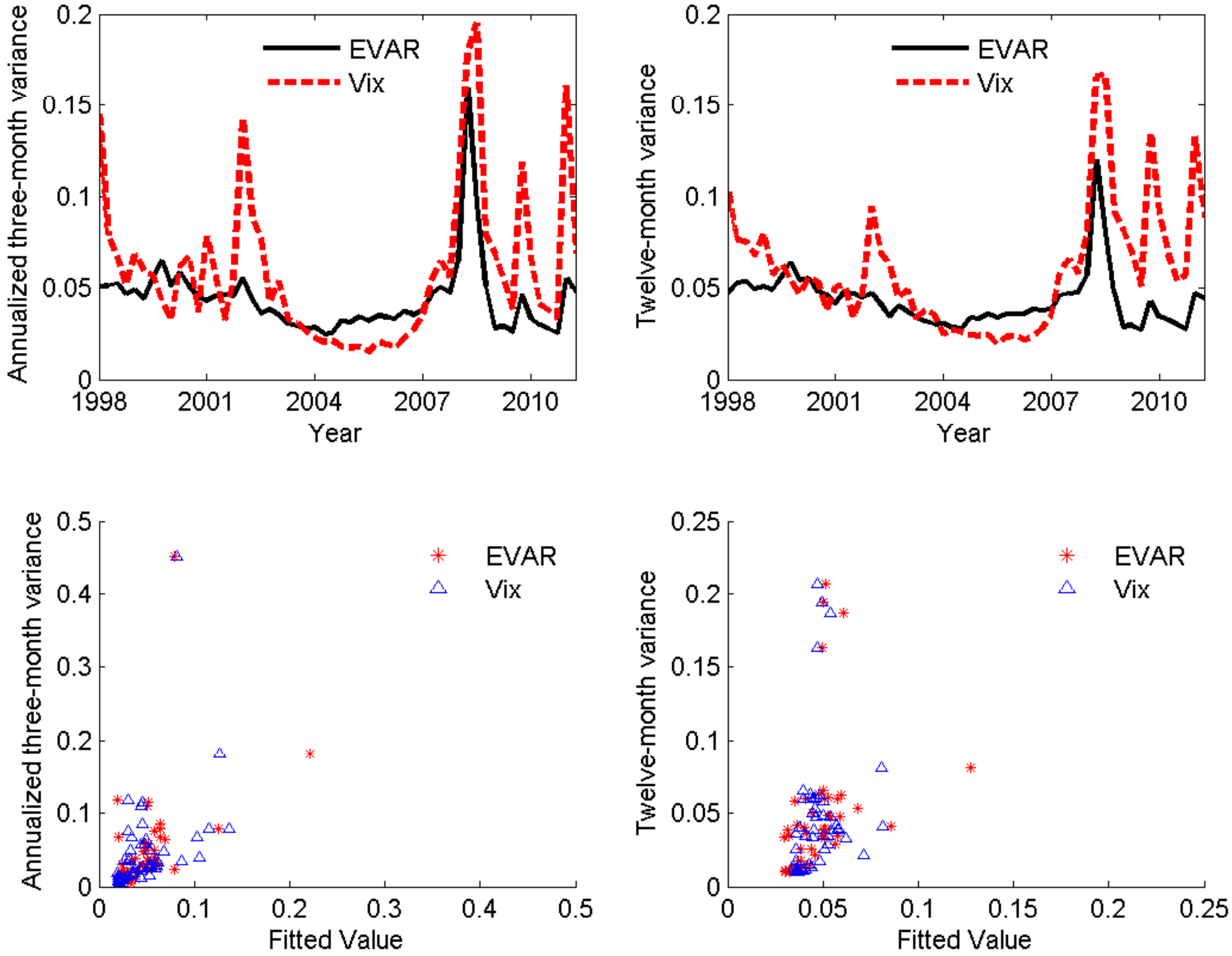

Figure 5: The top two diagrams correspond to forecasts of three-month (top left panel) and twelve-month (top right panel) variance from the VAR (VAR_h, solid black line) and from the option market (VIX_h^2, dashed red line). The bottom two diagrams correspond to scatter plots of three-month (bottom left panel) and twelve-month (bottom right panel) ralized variance against the corresponding forecast from the VAR (VAR_h, solid black line) and from the option market (VIX_h^2, dashed red line).

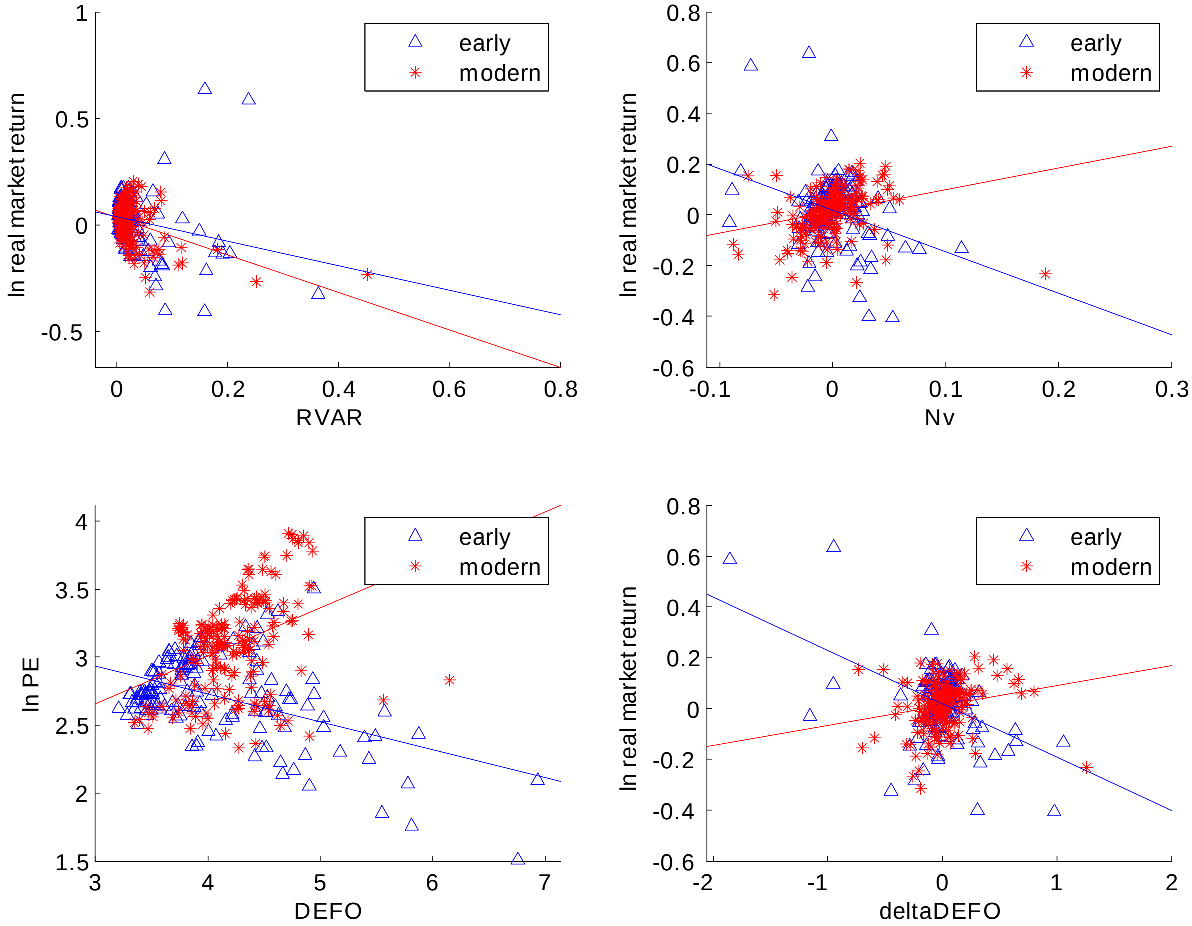

Figure 6: The top left portion of the figure plots the market return against $RVAR$. The top right portion of the figure plots the market return against volatility news, N_V. The bottom left of the figure plots PE against $DEFO$ (DEF orthogonalized to PE). The bottom right of the figure plots market returns against the contemporaneous change in $DEFO$, our simple proxy for news about long-horizon variance. In all four subplots, observations from the early period as denoted with blue triangles while observations from the modern period data are denoted with red asterisks.

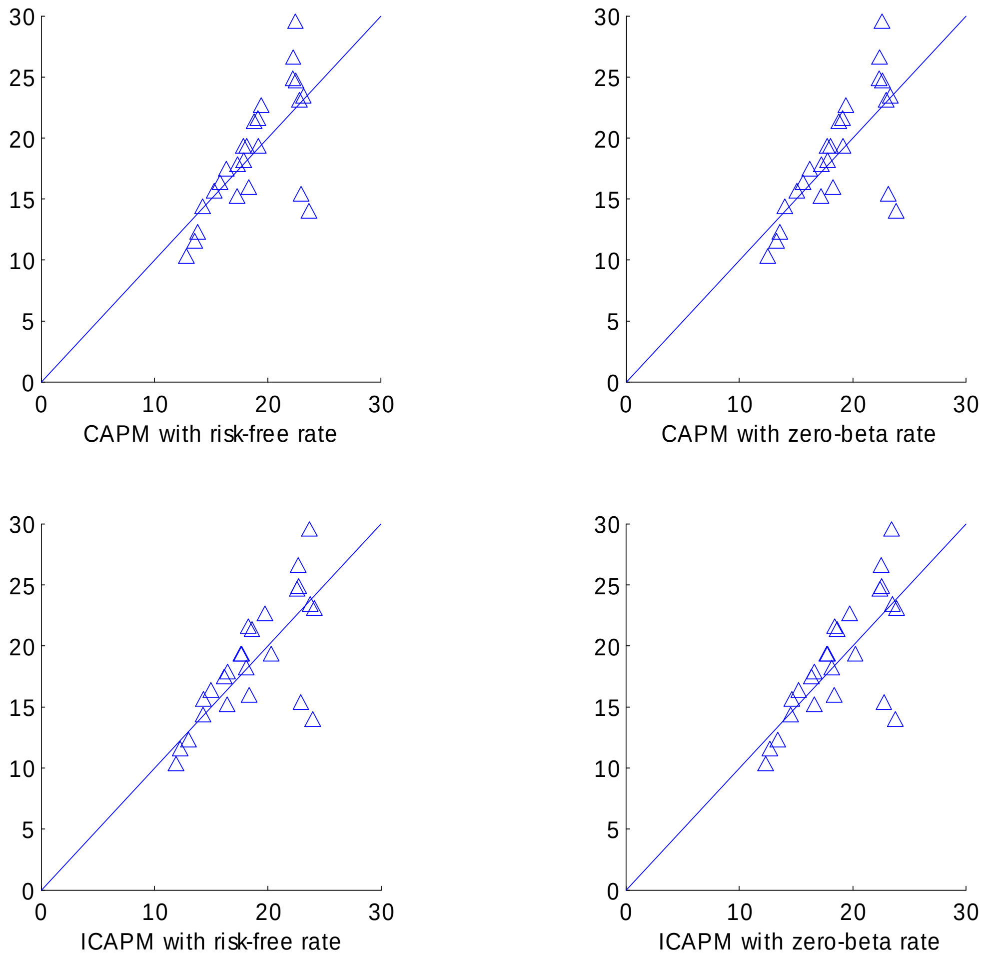

Figure 7: The four diagrams correspond to (clockwise from the top left) the CAPM with a constrained zero-beta rate, the CAPM with an unconstrained zero-beta rate, the three-factor ICAPM with a free zero-beta rate, and the three-factor ICAPM with the zero-beta rate constrained to the risk-freee rate. The horizontal axes correspond to the predicted average excess returns and the vertical axes to the sample average realized excess returns for the 25 ME- and BE/ME-sorted portfolios. The predicted values are from regressions presented in Table 7 for the sample period 1931:3-1963:2.

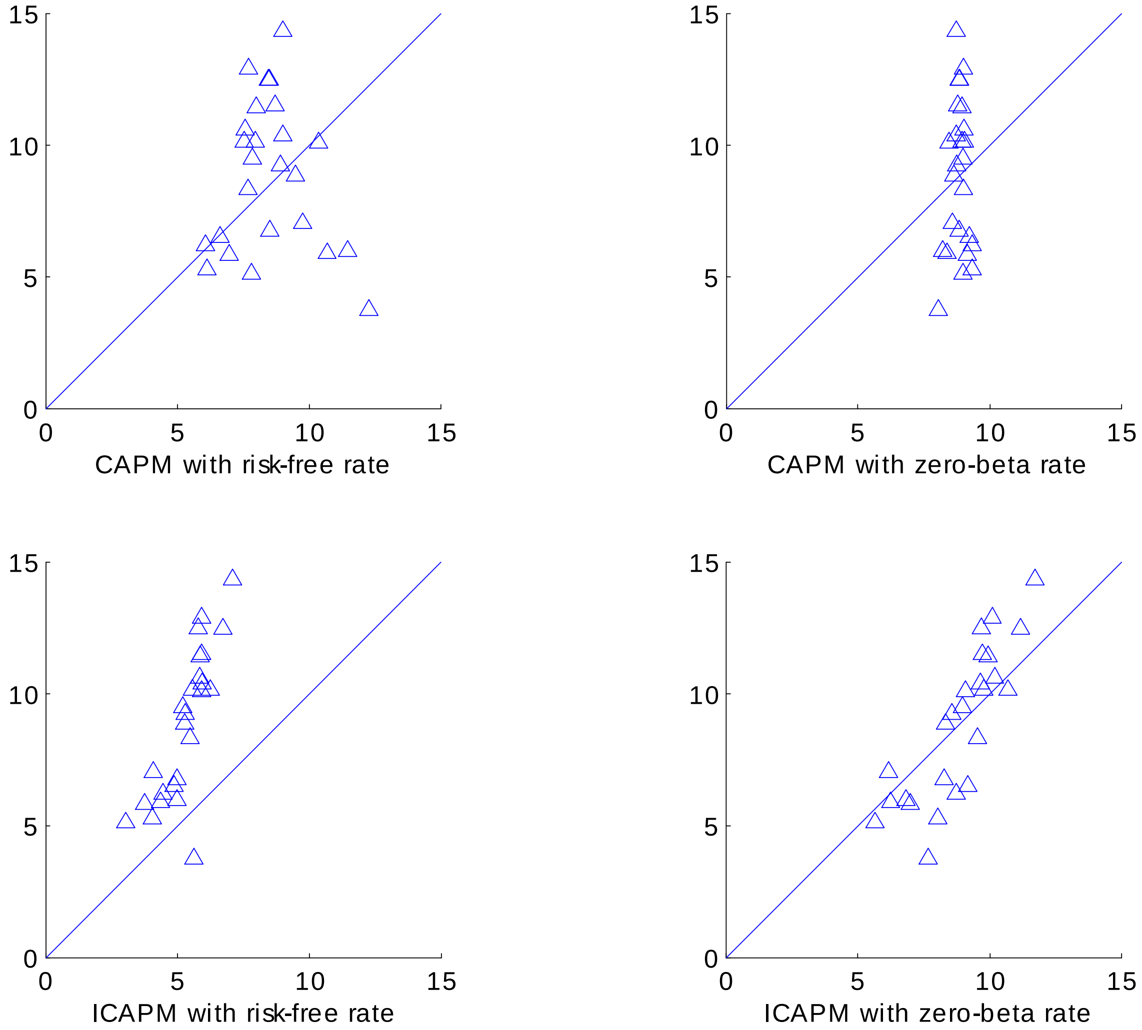

Figure 8: The four diagrams correspond to (clockwise from the top left) the CAPM with a constrained zero-beta rate, the CAPM with an unconstrained zero-beta rate, the three-factor ICAPM with a free zero-beta rate, and the three-factor ICAPM with the zero-beta rate constrained to the risk-freee rate. The horizontal axes correspond to the predicted average excess returns and the vertical axes to the sample average realized excess returns for the 25 ME- and BE/ME-sorted portfolios. The predicted values are from regressions presented in Table 8 for the sample period 1963:3-2011:4.

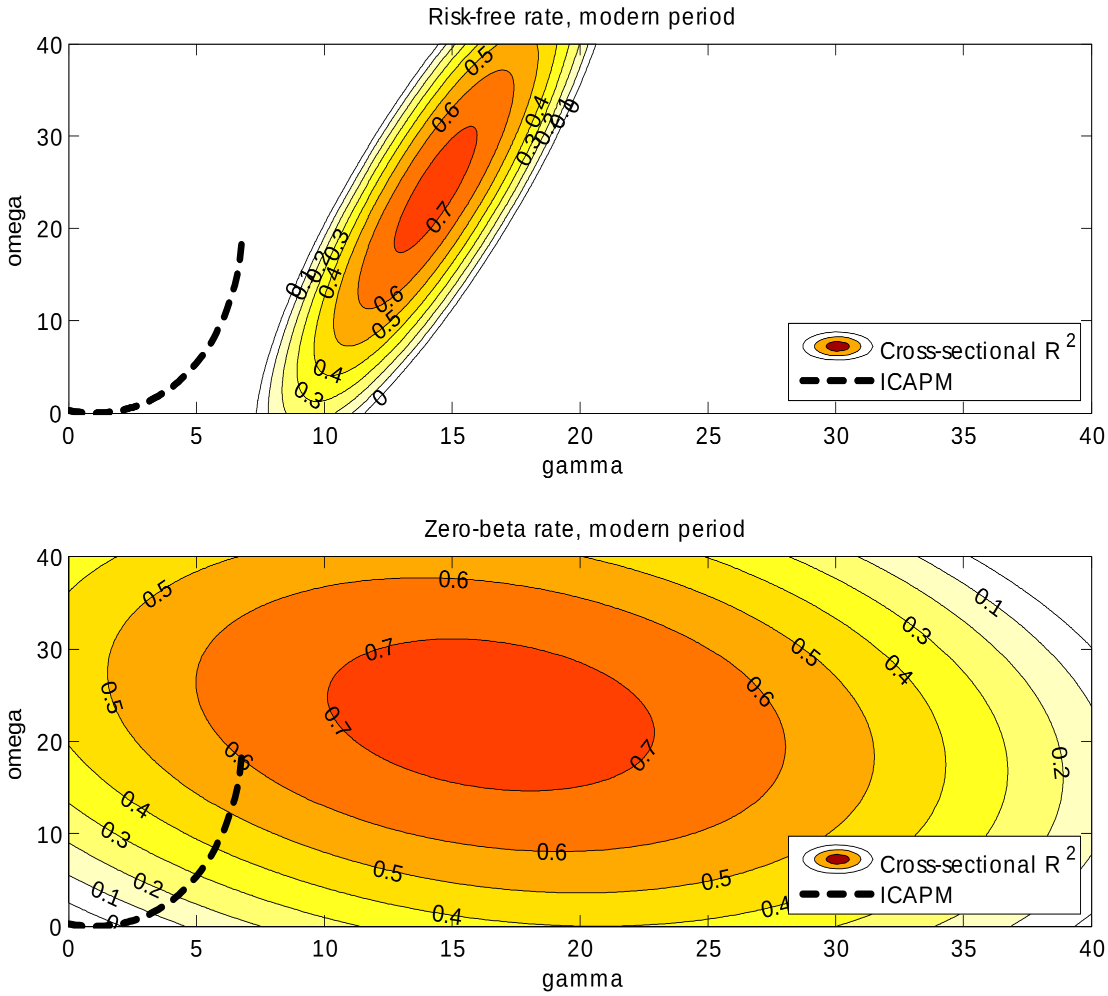

Figure 9: The two contour plots show how the R^2 of the cross-sectional regression explaining the average returns on the 25 size- and book-to-market portfolios varies for different values of γ and ω for the risk-free rate (top panel) and zero-beta rate (bottom panel) three-beta ICAPM model estimated in Table 8 for the sample period 1963:3-2011:4. The two plots also indicate the ICAPM relation between γ and ω described in equation (24).

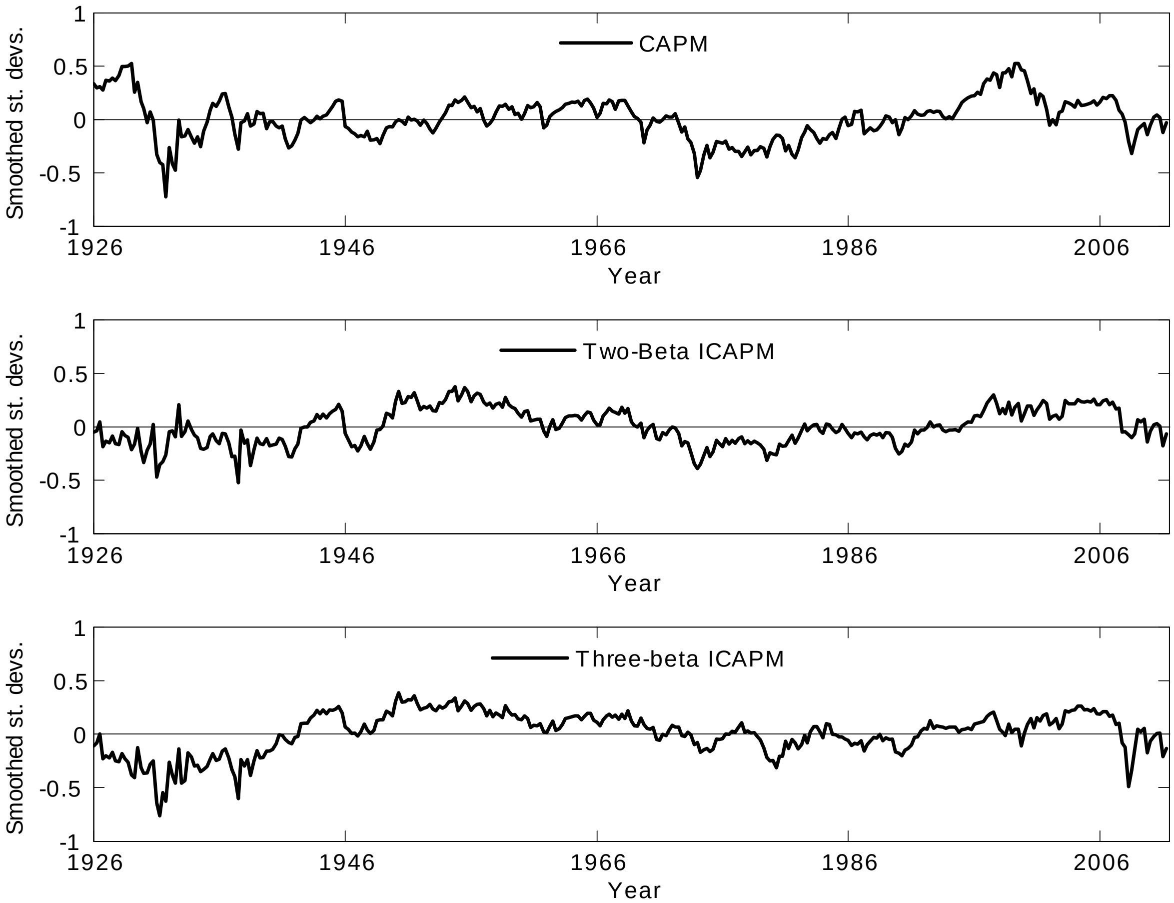

Figure 10: This figure plots the time-series of the smoothed combined shock for the CAPM ($N_{CF} - N_{DR}$), the two-beta ICAPM ($\gamma N_{CF} - N_{DR}$), and the three-beta ICAPM that includes stochastic volatility ($\gamma N_{CF} - N_{DR} - \frac{1}{2}\omega N_V$) for the unconstrained zero-beta rate specifications estimated in Table 8 for the modern subperiod. The shock is smoothed with a trailing exponentially-weighted moving average. The decay parameter is set to 0.08 per quarter, and the smoothed news series is generated as $MA_t(SDF) = 0.08 SDF_t + (1 - 0.08) MA_{t-1}(N)$. This decay parameter implies a half-life of six years. The sample period is 1926:2-2011:4.