Contents

List of Tables

Appendix to An Intertemporal CAPM with Stochastic Volatility

First draft: October 2011
This Version: January 2017

This Appendix provides a variety of supplemental information for "An Intertemporal CAPM with Stochastic Volatility"(ICSV).

1 Additional Literature Review

Our model is an example of an affine stochastic volatility model. 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. Similar models have been applied in the long-run risk literature by Eraker (2008), Eraker and Shaliastovich (2008), and Hansen (2012), but much of this literature uses volatility specifications that are not guaranteed to remain positive.

Two precursors to our work are unpublished papers by Chen (2003) and Sohn (2010).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.

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.

2 Detailed Model Derivations

In this section we derive an expression for the log stochastic discount factor (SDF) of the intertemporal CAPM model, and the corresponding pricing equations, when we allow for stochastic volatility. The SDF is based on Epstein–Zin utility, but imposes additional assumptions that allow us to express the SDF as a function of news about future cash flows, discount rates, and volatility, and obtain empirically testable implications.

2.1 The stochastic discount factor

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

The corresponding stochastic discount factor can be written as

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

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

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

The log SDF is a function of 1) consumption growth Δc_{t+1}, and 2) the log return on wealth r_{t+1}. In the remainder of this section, we show how to re-express the log SDF substituting consumption out, in a manner analogous to Campbell (1993) but allowing explicitly for time-varying volatility. We then discuss the implications of the model and its testable restrictions.

2.1.2 First step: 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 convenient identity (5) can therefore be used to write the log SDF (3) without reference to consumption growth:

$$m_{t+1} = \theta \ln \delta - \frac{\theta}{\psi} z_t + \frac{\theta}{\psi} h_{t+1} - \gamma r_{t+1}. \tag{6}$$

Given that the focus of our paper will be cross-sectional risk premia, it is useful to write the one-period innovation in the SDF:

$$m_{t+1} - \mathrm{E}_t m_{t+1} = \frac{\theta}{\psi}\left[h_{t+1} - \mathrm{E}_t h_{t+1}\right] - \gamma\left[r_{t+1} - \mathrm{E}_t r_{t+1}\right]. \tag{7}$$

As noted in Campbell (1993), consumption growth does not appear in this expression for the log SDF. Instead, the equation illustrates the dependence of the innovations in the SDF (which determine risk premia) on the one-period innovations in the wealth-consumption ratio and on the log return on the wealth portfolio. Next, we impose the asset pricing equation for the wealth portfolio and re-express the innovations in the SDF as a function of news about future cash flows, discount rates, and risk.

2.1.3 Second step: imposing the general pricing equation and lognormality to solve the SDF forward

We now add the assumption that asset returns and all state variables in the model are jointly conditionally lognormal. Since we allow for changing conditional volatility, we are careful to write second moments with time subscripts to indicate that they can vary over time. Under this standard assumption, the return on the wealth portfolio must satisfy:

$$0 = \ln \mathrm{E}_t \exp\{m_{t+1} + r_{t+1}\} = \mathrm{E}_t \left[m_{t+1} + r_{t+1}\right] + \frac{1}{2}\mathrm{Var}_t \left[m_{t+1} + r_{t+1}\right], \tag{8}$$

We can then substitute our log SDF (6) into the asset pricing equation (8) and multiply by $\frac{\psi}{\theta}$ to find an equation for z_t :

$$z_t = \psi \ln \delta + (\psi - 1)\mathrm{E}_t r_{t+1} + \mathrm{E}_t h_{t+1} + \frac{\psi}{\theta}\frac{1}{2}\mathrm{Var}_t \left[m_{t+1} + r_{t+1}\right]. \tag{9}$$

Next, 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 $\rho = \exp(\bar{z})/(1+\exp(\bar{z})) \approx 1 - C/W$. The two variables h_{t+1} and z_{t+1} are closely related: the former is the log ratio of wealth to consumption, $log(W_{t+1}/C_{t+1})$, the latter is the ratio of reinvested wealth to consumption, $log((W_{t+1} - C_{t+1})/C_{t+1})$. In fact, when the EIS, ψ, is 1, the loglinear relationship between the two variables holds exactly.

Combining the two equations (9) and (10) we then obtain an expression for the innovation in h_{t+1}:

$$\begin{aligned} h_{t+1} - \mathrm{E}_t h_{t+1} &= \rho(z_{t+1} - \mathrm{E}_t z_{t+1}) \\ &= (\mathrm{E}_{t+1} - \mathrm{E}_t)\rho\left((\psi - 1)r_{t+2} + h_{t+2} + \frac{\psi}{\theta}\frac{1}{2}\mathrm{Var}_{t+1}\left[m_{t+2} + r_{t+2}\right]\right). \end{aligned} \tag{11}$$

Solving forward to an infinite horizon,

$$\begin{aligned} h_{t+1} - \mathrm{E}_t h_{t+1} &= (\psi - 1)(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty} \rho^j r_{t+1+j} \\ &\quad + \frac{1}{2}\frac{\psi}{\theta}(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty} \rho^j \mathrm{Var}_{t+j}\left[m_{t+1+j} + r_{t+1+j}\right] \\ &= (\psi - 1)N_{DR,t+1} + \frac{1}{2}\frac{\psi}{\theta}N_{RISK,t+1}. \end{aligned} \tag{12}$$

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 equation for the innovations in the log SDF (7), and simplify to obtain:

$$
\begin{aligned}
m_{t+1} - \mathrm{E}_t m_{t+1} &= -\gamma \left[r_{t+1} - \mathrm{E}_t r_{t+1} \right] - (\gamma - 1) N_{DR,t+1} + \frac{1}{2} N_{RISK,t+1} \\
&= -\gamma N_{CF,t+1} - \left[-N_{DR,t+1} \right] + \frac{1}{2} N_{RISK,t+1} \qquad (13)
\end{aligned}
$$

Equation (13) expresses the log SDF in terms of the market return and news about future variables. In particular, it identifies three priced factors: the market return (with a price of risk γ), discount rate news (with price of risk $(\gamma - 1)$), and news about future risk (with price of risk of $-\frac{1}{2}$). This is an extension of the ICAPM as derived by Campbell (1993), with no reference to consumption or the elasticity of intertemporal substitution ψ. When the investor's risk aversion is greater than 1, assets which hedge aggregate discount rates (negative covariance with N_{DR}) or aggregate risk (positive covariance with N_{CF}) will have lower expected returns, all else equal.

The second equation rewrites the model, following Campbell and Vuolteenaho (2004), by breaking the market return into cash-flow news and discount-rate news. Cash-flow news $N_{CF,t+1}$ is defined by $N_{CF,t+1} = r_{t+1} - \mathrm{E}_t r_{t+1} + N_{DR,t+1}$. The price of risk for cash-flow news is γ times greater than the 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 (13) shows the risk price for exposure to news about future risks and did not appear in Campbell and Vuolteenaho's model, which assumed homoskedasticity. Not surprisingly, the coefficient is positive, indicating that an asset providing positive returns when risk expectations increase will offer a lower return on average (the log SDF is high when future volatility is anticipated to be high).

While the elasticity of intertemporal substitution ψ does not affect risk prices (and therefore risk premia) in our model, this parameter does influence the implied behavior of the investor's consumption.

2.1.4 Third step: linking news about risk to news about volatility

The risk news term $N_{RISK,t+1}$ in equation (13) represents news about the conditional volatility of returns plus the stochastic discount factor, $\mathrm{Var}_t \left[m_{t+1} + r_{t+1} \right]$. It therefore depends on the SDF m and its innovations. To close the model and derive its empirical implications, we need to add assumptions on the data generating process for stock returns and the variance terms that will allow to solve for the term $\text{Var}_t [m_{t+1} + r_{t+1}]$ and compute the news terms. These assumptions will imply that the conditional volatility of returns plus the stochastic discount factor is proportional to the conditional volatility of returns themselves.

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

$$\mathbf{x}_{t+1} = \bar{\mathbf{x}} + \mathbf{\Gamma} (\mathbf{x}_t - \bar{\mathbf{x}}) + \sigma_t \mathbf{u}_{t+1}, \tag{14}$$

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

$$\begin{aligned} N_{DR,t+1} &= (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j} \\ &= \mathbf{e}_1' \sum_{j=1}^{\infty} \rho^j \mathbf{\Gamma}^j \sigma_t \mathbf{u}_{t+1} \\ &= \mathbf{e}_1' \rho \mathbf{\Gamma} (\mathbf{I} - \rho \mathbf{\Gamma})^{-1} \sigma_t \mathbf{u}_{t+1}, \end{aligned} \tag{15}$$

while implied cash flow news is:

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

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, $\text{Var}_t [m_{t+1} + r_{t+1}] \propto \sigma^2_t$. As a result, the conditional variance of the scaled variables, $\text{Var}_t [(m_{t+1} + r_{t+1}) / \sigma_t] = \omega_t$, will be a constant that does not depend on the state variables: ω. Without knowing the parameters of the utility function, we can write $\text{Var}_t [m_{t+1} + r_{t+1}] = \omega \sigma^2_t$, so that the news about risk, N_{RISK}, is proportional to news about market return variance, N_V.

$$\begin{aligned} N_{RISK,t+1} &= (E_{t+1}-E_t)\sum_{j=1}^{\infty}\rho^j \mathrm{Var}_{t+j}\left[r_{t+1+j}+m_{t+1+j}\right] \\ &= (E_{t+1}-E_t)\sum_{j=1}^{\infty}\rho^j\left(\omega\sigma_{t+j}^2\right) \\ &= \omega\rho\mathbf{e}_2'\sum_{j=0}^{\infty}\rho^j\mathbf{\Gamma}^j\sigma_t\mathbf{u}_{t+1} \\ &= \omega\rho\mathbf{e}_2'\left(\mathbf{I}-\rho\mathbf{\Gamma}\right)^{-1}\sigma_t\mathbf{u}_{t+1}=\omega N_{V,t+1}. \end{aligned} \tag{17}$$

2.2 Solving for ω

We now show how to solve for the unknown parameter ω. From the definition of ω,

$$\begin{aligned} \omega\sigma_t^2 &= \mathrm{Var}_t\left[m_{t+1}+r_{t+1}\right] \\ &= \mathrm{Var}_t\left[\frac{\theta}{\psi}h_{t+1}+(1-\gamma)r_{t+1}\right] \\ &= \mathrm{Var}_t\left[\frac{\theta}{\psi}\left((\psi-1)N_{DR,t+1}+\frac{1}{2}\frac{\psi}{\theta}\omega N_{V,t+1}\right)+(1-\gamma)r_{t+1}\right] \\ &= \mathrm{Var}_t\left[(1-\gamma)N_{DR,t+1}+\frac{1}{2}\omega N_{V,t+1}+(1-\gamma)r_{t+1}\right] \\ &= \mathrm{Var}_t\left[(1-\gamma)N_{CF,t+1}+\frac{1}{2}\omega N_{V,t+1}\right] \\ &= (1-\gamma)^2\mathrm{Var}_t\left[N_{CF_{t+1}}\right]+\omega(1-\gamma)\mathrm{Cov}_t\left[N_{CF_{t+1}},N_{V_{t+1},}\right]+\frac{\omega^2}{4}\mathrm{Var}_t\left[N_{V_{t+1}}\right]. \end{aligned} \tag{18}$$

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:

$$\frac{1}{\sigma_t}N_{CF,t+1} = \mathbf{x}_{CF}\mathbf{u}_{t+1}=\left(\mathbf{e}_1'+\mathbf{e}_1'\rho\mathbf{\Gamma}(\mathbf{I}-\rho\mathbf{\Gamma})^{-1}\right)\mathbf{u}_{t+1} \tag{19}$$

$$\frac{1}{\sigma_t}N_{V,t+1} = \mathbf{x}_V\mathbf{u}_{t+1}=\left(\mathbf{e}_2'\rho(\mathbf{I}-\rho\mathbf{\Gamma})^{-1}\right)\mathbf{u}_{t+1}. \tag{20}$$

Then ω solves

$$0=\omega^2\frac{1}{4}\mathbf{x}_V\mathbf{\Sigma}\mathbf{x}_V'-\omega\left(1-(1-\gamma)\mathbf{x}_{CF}\mathbf{\Sigma}\mathbf{x}_V'\right)+(1-\gamma)^2\mathbf{x}_{CF}\mathbf{\Sigma}\mathbf{x}_{CF}' \tag{21}$$

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 current risk through future risk: ω appears on the right-hand side of the equation as well. Given that in our estimation we find $\text{Cov}_t\left[N_{CF,t+1}, N_{V,t+1}\right] < 0$, this second effect makes ω increase even faster with γ.

2.2.1 Selecting the correct root of the quadratic equation

The equation defining ω will generally have two solutions

$$\omega = \frac{1-(1-\gamma)\, x_{CF}\mathbf{\Sigma}x_V' \pm \sqrt{\left(1-(1-\gamma)\, x_{CF}\mathbf{\Sigma}x_V'\right)^2 - (1-\gamma)^2\left(x_V\mathbf{\Sigma}x_V'\right)\left(x_{CF}\mathbf{\Sigma}x_{CF}'\right)}}{\frac{1}{2}x_V\mathbf{\Sigma}x_V'}. \tag{22}$$

While the (approximate) Euler equation holds for both solutions, the correct solution is the one with the negative sign on the radical. This result can be confirmed from numerical computation, and it can also be easily seen by observing the behavior of the solutions in the limit as volatility news goes to zero and the model become homoskedastic. With the false solution, ω becomes infinitely large as $x_V \to 0$. This false solution corresponds to the log value of invested wealth going to negative infinity. On the other hand, we can exploit that the correct solution for ω converges to $(1-\gamma)^2 x_{CF}\mathbf{\Sigma}x_{CF}'$. This is what we would expect, since in that case $\omega = \text{Var}_t\left[(1-\gamma)\, N_{CF,t+1}/\sigma_t\right]$.

2.2.2 Simplifying the existence condition for a real root

Appendix Figure 1 plots ω as a function of γ, conditional on our VAR parameter estimates. The upper bound of 7.2 for γ is the value of γ above which a real solution to the quadratic equation ceases to exist.

The existence condition for a solution for ω corresponds to the following inequality:

$$\left[1-(1-\gamma)(x_{CF}'\Sigma x_V)\right]^2 - (1-\gamma)^2\left(x_V'\Sigma x_V\right)\left(x_{CF}'\Sigma x_{CF}\right) \geq 0 \tag{23}$$

We show here that this condition can be simplified to a set of bounds on γ of the form:

$$1 - \frac{1}{(\rho_n+1)\sigma_{cf}\sigma_v} \leq \gamma \leq 1 - \frac{1}{(\rho_n-1)\sigma_{cf}\sigma_v} \tag{24}$$

where ρ_n is the correlation of the news terms, σ_{cf} is the scaled standard deviation of cash flow news, and σ_v is the scaled standard deviation of volatility news. Note that since $-1 \leq \rho_n \leq 1$, the lower bound on γ is always (weakly) below 1, and the upper bound is always (weakly) above 1. We also note that empirically, the lower bound is often below zero, and therefore not actually binding. For example, in our case depicted in Appendix Figure 1, only the upper bound on γ is binding, as the lower bound from equation (24) lies below zero.

As evident from equation (23), the existence condition is itself a simple quadratic inequality in $(1-\gamma)$. We can rewrite it as:

$$(1-\gamma)^2(x'_{CF}\Sigma x_V)^2 + 1 - 2(1-\gamma)(x'_{CF}\Sigma x_V) - (1-\gamma)^2\left(x'_{CF}\Sigma x_{CF}\right)\left(x'_V\Sigma x_V\right) \geq 0 \tag{25}$$

or:

$$(1-\gamma)^2\left[(x'_{CF}\Sigma x_V)^2 - \left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)\right] - 2(1-\gamma)(x'_{CF}\Sigma x_V) + 1 \geq 0 \tag{26}$$

The two roots of this equation can be found as:

$$\begin{aligned}(1-\gamma) &= \frac{2(x'_{CF}\Sigma x_V) \pm \sqrt{4(x'_{CF}\Sigma x_V)^2 - 4\left[(x'_{CF}\Sigma x_V)^2 - \left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)\right]}}{2\left[(x'_{CF}\Sigma x_V)^2 - \left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)\right]} \\ &= \frac{(x'_{CF}\Sigma x_V) \pm \sqrt{\left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)}}{\left[(x'_{CF}\Sigma x_V)^2 - \left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)\right]}\end{aligned} \tag{27}$$

Note that this equation always has two real solutions, since $\left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right) > 0$. The denominator can be written as:

$$\begin{aligned}&\left[(x'_{CF}\Sigma x_V)^2 - \left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)\right] = \left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)\left(\rho_n^2 - 1\right) \\ &= \left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)(\rho_n+1)(\rho_n-1) = \sigma_v^2\sigma_{cf}^2(\rho_n+1)(\rho_n-1)\end{aligned} \tag{28}$$

while the numerator can be written as:

$$(x'_{CF}\Sigma x_V) \pm \sqrt{\left(x'_V\Sigma x_V\right)\left(x'_{CF}\Sigma x_{CF}\right)} = \sigma_v\sigma_{cf}\rho_n \pm \sigma_v\sigma_{cf} = \sigma_v\sigma_{cf}(\rho_n \pm 1) \tag{29}$$

Therefore, the two roots can be found as:

$$(1-\gamma) = \frac{\sigma_v\sigma_{cf}(\rho_n \pm 1)}{\sigma_v^2\sigma_{cf}^2(\rho_n+1)(\rho_n-1)} = \frac{(\rho_n \pm 1)}{\sigma_v\sigma_{cf}(\rho_n+1)(\rho_n-1)} \tag{30}$$

Or:

$$\overline{(1-\gamma)} = \frac{(\rho_n - 1)}{\sigma_v\sigma_{cf}(\rho_n+1)(\rho_n-1)} = \frac{1}{\sigma_v\sigma_{cf}(\rho_n+1)} \geq 0 \tag{31}$$

and

$$\underline{(1-\gamma)} = \frac{(\rho_n + 1)}{\sigma_v\sigma_{cf}(\rho_n+1)(\rho_n-1)} = \frac{1}{\sigma_v\sigma_{cf}(\rho_n-1)} \leq 0 \tag{32}$$

Finally, we note that since $[(x'_{CF}\Sigma x_V)^2 - (x'_V\Sigma x_V)\,(x'_{CF}\Sigma x_{CF})] = \sigma_v^2\sigma_{cf}^2(\rho_n+1)(\rho_n-1) \leq 0$, the quadratic inequality (23) will have solutions between the two roots, for

$$1-\gamma \leq \overline{(1-\gamma)} = \frac{1}{\sigma_v\sigma_{cf}(\rho_n+1)} \tag{33}$$

and for

$$1-\gamma \geq \underline{(1-\gamma)} = \frac{1}{\sigma_v\sigma_{cf}(\rho_n-1)} \tag{34}$$

or equivalently:

$$\gamma \geq 1 - \frac{1}{\sigma_v\sigma_{cf}(\rho_n+1)} \tag{35}$$

$$\gamma \leq 1 - \frac{1}{\sigma_v\sigma_{cf}(\rho_n-1)} \tag{36}$$

2.3 Derivation of the moment conditions

After solving for ω, we can rewrite the stochastic discount factor as:

$$m_{t+1} - \mathrm{E}_t m_{t+1} = -\gamma N_{CF,t+1} - [-N_{DR,t+1}] + \frac{1}{2}\omega N_{V,t+1} \tag{37}$$

To derive the moment conditions of the model, we go back to the general asset pricing equation under lognormality

$$0 = \ln \mathrm{E}_t \exp\{m_{t+1} + r_{i,t+1}\} = \mathrm{E}_t\left[m_{t+1} + r_{i,t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[m_{t+1} + r_{i,t+1}\right]. \tag{38}$$

The same equation can be rewritten as:

$$0 = \mathrm{E}_t\left[m_{t+1}\right] + \mathrm{E}_t\left[r_{i,t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[m_{t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[r_{i,t+1}\right] + Cov_t(r_{i,t+1}, m_{t+1} - \mathrm{E}_t m_{t+1}) \tag{39}$$

We now rearrange this equation and make two substitutions. First, we note that the conditional mean of the log SDF innovation is zero, so that

$$Cov_t(r_{i,t+1}, m_{t+1} - \mathrm{E}_t m_{t+1}) = \mathrm{E}_t\left[r_{i,t+1}(m_{t+1} - \mathrm{E}_t m_{t+1})\right] \tag{40}$$

Second, we note that

$$E_t r_{i,t+1} + \frac{1}{2}\sigma_{it}^2 \simeq (E_t R_{i,t+1} - 1) \tag{41}$$

which links the expected log returns (adjusted by their variance) to the expected gross level returns $r_{i,t+1}$.[1]

After these two substitutions, we can rearrange (39) to yield:

$$E_t R_{i,t+1} - 1 = -\mathrm{E}_t\left[m_{t+1}\right] - \frac{1}{2}\mathrm{Var}_t\left[m_{t+1}\right] - \mathrm{E}_t\left[r_{i,t+1}(m_{t+1} - \mathrm{E}_t m_{t+1})\right] \tag{42}$$

Given any reference asset j (which could be but does not need to be the risk-free rate), we can write the relative risk premium of i relative to j as:

$$E_t\left[R_{i,t+1} - R_{j,t+1}\right] = -\mathrm{E}_t\left[(r_{i,t+1} - r_{j,t+1})(m_{t+1} - \mathrm{E}_t m_{t+1})\right] \tag{43}$$

by taking the difference of equation (42) between i and j. We can then substitute the expression for the innovations in the SDF and write:

$$E_t\left[R_{i,t+1} - R_{j,t+1}\right] = \mathrm{E}_t\left[(r_{i,t+1} - r_{j,t+1})(\gamma N_{CF,t+1} + [-N_{DR,t+1}] - \frac{1}{2}\omega N_{V,t+1})\right] \tag{44}$$

2.4 A simple fully-solved example with $\psi = 1$

In this section we solve analytically for a simple model with $\psi = 1$ and $\gamma > 1$. We show that for the value function to exist the parameters of the model must satisfy a quadratic equation, and we show that in this model the equation corresponds to equation (12) in ICSV (2016), i.e. the upper bound on γ that ensures existence of a real solution for the price of volatility risk ω (only the upper bound matters here, since we are looking at the case $\gamma > 1$). For tractability purposes, we assume that consumption growth is *iid*, so the only state variable will be volatility. Finally, we consider separately the existence conditions for the case of a homoskedastic volatility process.

Since $\psi = 1$, we can write the log value function relative to consumption, $v_t = \ln(V_t/C_t)$, recursively as (see Hansen, Heaton and Li 2008):

$$v_t = \frac{\delta}{1-\gamma}\ln \mathrm{E}_t \exp\left\{(1-\gamma)(v_{t+1} + \Delta c_{t+1})\right\} \tag{45}$$

Assume that volatility and consumption growth follow the process

$$\sigma_{t+1}^2 = s + d\sigma_t^2 + x\sigma_t\epsilon_{t+1} \tag{46}$$

$$\Delta c_{t+1} = k\sigma_t\eta_{t+1} \tag{47}$$

with ϵ_t and η_t normal with unit standard deviation, and correlation θ. d captures the

---

[1] By lognormality, we have: $E_t[r_{i,t+1}] + \frac{\sigma_{i,t}^2}{2} = lnE_t[R_{i,t+1}]$. Now, for the expected gross return $E_t[R_{i,t+1}]$ close to 1, we will have: $lnE_t[R_{i,t+1}] \simeq E_t[R_{i,t+1}] - 1$, from which the result follows.

persistence of volatility, while x scales the volatility of volatility.

2.4.1 Existence of a solution in the simple model

We conjecture that v and Δc are jointly lognormal, and write:

$$\begin{aligned} v_t &= \frac{\delta}{1-\gamma} \ln \mathrm{E}_t \exp\left\{(1-\gamma)(v_{t+1}+\Delta c_{t+1})\right\} \\ &= \frac{\delta}{1-\gamma} \left[\mathrm{E}_t \left\{(1-\gamma)(v_{t+1}+\Delta c_{t+1})\right\} + 0.5\mathrm{Var}_t \left\{(1-\gamma)(v_{t+1}+\Delta c_{t+1})\right\}\right] \\ &= \delta \mathrm{E}_t \left\{v_{t+1}+\Delta c_{t+1}\right\} + \delta(1-\gamma)0.5\mathrm{Var}_t \left\{v_{t+1}+\Delta c_{t+1}\right\}. \end{aligned} \tag{48}$$

Since consumption has mean zero in the simple model,

$$v_t = \delta \mathrm{E}_t \left\{v_{t+1}\right\} + \delta(1-\gamma)0.5\mathrm{Var}_t \left\{v_{t+1}+\Delta c_{t+1}\right\} \tag{49}$$

We now guess that the log value function is linear in σ_t^2:

$$v_t = a + b\sigma_t^2 \tag{50}$$

and obtain:

$$\begin{aligned} a + b\sigma_t^2 &= \delta(a + b\mathrm{E}_t \left\{\sigma_{t+1}^2\right\}) + \delta(1-\gamma)0.5\mathrm{Var}_t \left\{a + b\sigma_{t+1}^2 + k\sigma_t\eta_{t+1}\right\} \\ &= \delta(a + b\mathrm{E}_t \left\{s + d\sigma_t^2\right\}) + \delta(1-\gamma)0.5\mathrm{Var}_t \left\{a + b\sigma_{t+1}^2 + k\sigma_t\eta_{t+1}\right\} \\ &= \delta(a + b\mathrm{E}_t \left\{s + d\sigma_t^2\right\}) + \delta(1-\gamma)0.5\mathrm{Var}_t \left\{bx\sigma_t\epsilon_{t+1} + k\sigma_t\eta_{t+1}\right\} \\ &= \delta(a + bs + bd\sigma_t^2) + \delta(1-\gamma)0.5[b^2x^2 + k^2 + 2bxk\theta]\sigma_t^2. \end{aligned} \tag{51}$$

Matching coefficients on σ_t^2:

$$b = \delta bd + \delta(1-\gamma)0.5(b^2x^2 + k^2 + 2bxk\theta) \tag{52}$$

or:

$$\left(\delta(1-\gamma)0.5x^2\right) b^2 + \left(\delta d - 1 + \delta(1-\gamma)xk\theta\right) b + \left(\delta(1-\gamma)0.5k^2\right) = 0 \tag{53}$$

This is a quadratic equation which may not have a solution. For the solution to exist, we need:

$$\left(\delta d - 1 + \delta(1-\gamma)xk\theta\right)^2 > \left(\delta(1-\gamma)x^2\right)\left(\delta(1-\gamma)k^2\right) \tag{54}$$

Given the signs of these variables, this equation can be rewritten as:

$$1 - \delta d - \delta(1-\gamma)xk\theta > \delta(\gamma-1)xk \tag{55}$$

Rearranging, the existence condition for the value function in this model is given by:

$$(\gamma - 1) \leq \frac{1 - \delta d}{k\delta x(1-\theta)}. \quad (56)$$

2.4.2 Comparison with the existence of a real solution to ω

We can compare this equation with the condition for having a real solution for the price of volatility risk ω in the general model of ICSV (2016). We can rewrite that upper bound on γ (from equation 24) as:

$$(\text{Corr}(N_{cf}, N_V) - 1)(1-\gamma)\sigma_{cf}\sigma_v \leq 1. \quad (57)$$

We now apply this condition to the fully solved model presented above. In this model we have:

$$N_{CF,t+1} = \Delta c_{t+1} = k\sigma_t \eta_{t+1} \quad (58)$$

$$\begin{aligned} N_{V,t+1} &= (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\delta^j \sigma^2_{t+j} = (\mathrm{E}_{t+1} - \mathrm{E}_t)\delta\sum_{j=0}^{\infty}\delta^j\sigma^2_{t+j+1} \\ &= \delta\sum_{j=0}^{\infty}\delta^j d^j(x\sigma_t\epsilon_{t+1}) = \frac{\delta}{1-\delta d}x\sigma_t\epsilon_{t+1}. \end{aligned} \quad (59)$$

$$\text{Corr}(N_{CF}, N_{DR}) = \theta \quad (60)$$

Substituting:

$$(\theta - 1)(1-\gamma)\frac{k\delta x}{1-\delta d} \leq 1 \quad (61)$$

or

$$(\gamma - 1) \leq \frac{1-\delta d}{k\delta x(1-\theta)} \quad (62)$$

which precisely coincides with the existence condition for v_t shown in the previous subsection.

3 VAR summary statistics

We report summary statistics for the variables in our VAR in Appendix Table 1. A comparison of the unscaled and scaled autocorrelation matrices, provided in Appendix Table 2, documents that much of the sample autocorrelation in the unscaled residuals is eliminated by our WLS approach.

4 Predicting Long-Run Volatility

The predictability of volatility, and especially of its long-run component, is central to this paper. In the text, 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 Appendix Table 3 we regress realized, discounted, annualized long-run variance up to period h,

$$LHRVAR_h = \frac{4\Sigma_{j=1}^{h} \rho^{j-1} RVAR_{t+j}}{\Sigma_{j=1}^{h} \rho^{j-1}}, \tag{63}$$

on both our VAR forecast and some alternative forecasts of long-run variance.[2] We focus our discussion on the 10-year horizon ($h = 40$) as longer horizons come at the cost of fewer independent observations; however, Appendix Table 4 confirms that our results are robust to horizons ranging from one to 15 years.

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 $LHRVAR_{40}$.[3] To these two models, we add the set of variables from our VAR, and compare the forecasting ability of these different models.

Appendix Table 3 Panel A reports the results of forecasting regressions of long-run volatil-

[2] 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.

[3] 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).

ity $LHRVAR_{40}$ using different specifications. The first regression presents results using the state variables in our VAR, each included separately. The second regression predicts $LHRVAR_{40}$ 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. Appendix Table 4 documents that these conclusions are true at all horizons (with the exception of $RVAR$ at $h = 8$ and $h = 20$, i.e. two and five years). Finally, the coefficient on the VAR-implied forecast, VAR_{40}, is 1.02. 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 essentially measured as zero. Appendix Table 4 shows that this finding is true at all horizons we consider.

The bottom part of Appendix Table 3 Panel A examines more carefully the link between DEF and $LHRVAR_{40}$. Regressions nine 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 nearly 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.

As a further check on the usefulness of our VAR approach, we compare our variance forecasts to option-implied variance forecasts. Specifically, using option data from OptionMetrics for the period 1998–2011, we construct the synthetic prices of variance swaps (claims to the realized variance from inception to the maturity of the contract), replicated using a portfolio of options. We construct these prices for maturities 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 all maturities. We 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^2_h).[4] Since our VAR is quarterly, we study forecasts at the three-month, six-month, nine-month, and twelve-month horizons. The top panels of Appendix Figure 2 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. Appendix Figure 2 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^2_h forecasts at the same horizon. Appendix Table 3 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.

Appendix Table 3 Panel C reports regressions forecasting $LHRVAR_h$ using the VAR_h forecast, the VIX^2_h 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 Appendix Figure 2 plot $LHRVAR_h$ against the fitted value from the VAR_h forecast and against the fitted value of the VIX^2_h forecast for the three-month and twelve-month horizons. The figure confirms that VAR_h is as informative as VIX^2_h, if not more so. Indeed, Appendix Table 3 Panel C shows that when both forecasts are included in the regression, VAR_h subsumes VIX^2_h, remaining statistically and economically significant.

Taken together, these results make a strong case that credit spreads and valuation ratios contain information about future volatility not captured by simple univariate models, even those like the FIGARCH model or the two-component EGARCH model that are designed to

[4]As the VIX^2_h 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 $LHRVAR_h$.

fit long-run movements in volatility, and that our VAR method for calculating long-horizon forecasts preserves this information.

4.1 Shocks to short- and long-run expected variance

Our empirical specification (the VAR that drives volatility and the other state variables) delivers a time-series model where volatility is affected by different shocks, each with different time-series properties. Some shocks affect volatility for a long time, some for a short time. Our multifactor specification predicts that one-period and multi-period innovations are not perfectly correlated. Our results are in stark contrast to a univariate model of volatility (e.g. if volatility followed an AR(1)). In that case, one-period shocks and longer-horizon shocks (like N_V) would be perfectly correlated. Appendix Table 5 summarizes this message: in the data, N_V and innovations in one-period volatility expectations are not perfectly correlated, as predicted in a multifactor world. Appendix Figure 3 makes this point graphically.

5 Comparison with BKSY: details

In this section we discuss details of the comparison with Bansal, Kiku, Shaliastovich, and Yaron's (BKSY 2014) model (Section 7.1 in ICSV). We review the theoretical implications of a homoskedastic volatility process, we compare by simulation the homoskedastic and heteroskedastic versions of the volatility process, and we present a comparison of the empirical results in our and BKSY's papers.

5.1 A Homoskedastic Stochastic Volatility Model

It is interesting to explore the alternative hypothesis of a homoskedastic process for σ_t^2 (as in BKSY). We show in the paper that under the assumption that σ_t^2 scales *all* the shocks of the VAR, we obtain the result that $\mathrm{Var}_t(RISK_{t+1}) \equiv \mathrm{Var}_t(m_{t+1} + r_{t+1}) = \omega\sigma_t^2$, so that $N_{RISK} = \omega N_V$. Given this proportionality, in our empirical analysis we can use N_V as a pricing factor, with a price of risk of ω. We now explore whether this proportionality holds under the assumption of homoskedasticity of the variance process.

For N_{RISK} to be proportional to N_V, a sufficient condition is that $\mathrm{Var}_t(RISK_{t+1})$ is proportional (as in our case) or at least affine (as in BKSY) in σ_t^2. If this is not the case, the news terms will not generally be proportional to each other, and it will not generally be appropriate to use N_V as a pricing factor.

When considering the homoskedastic volatility case, it is important to define which shocks are actually homoskedastic. When the volatility process σ_t^2 is modeled as part of a VAR, the fact that its own innovation has constant variance does *not* imply that N_V will also have constant variance. To see this, call η_{t+1} the *unscaled* vector of VAR innovations. If σ_t^2 is the $v-th$ element of the VAR and its shock has constant variance, but the other shocks are scaled by σ_t^2, then we will have: $\text{Var}_t(e_v'\eta_{t+1})$ equal to a constant but $\text{Var}_t(e_{i\neq v}'\eta_{t+1}) \propto \sigma_t^2$ (where e_i is a vector of zeros with 1 at the i-th element). Now consider that the volatility news term N_V can be expressed as $\lambda_v'\eta_{t+1}$, where $\lambda_v = e_v'^{-1}$. In general N_V will *not* have constant variance, but rather its variance will be a linear function of σ_t and σ_t^2. With a simple example, suppose that σ_t^2 is the second element of a 2-variable VAR, and $\lambda_v = [\lambda_1\, \lambda_2]'$. Then,

$$\text{Var}_t(N_{V,t+1}) = \text{Var}_t(\lambda_1 \sigma_t u_{1,t+1} + \lambda_2 u_{2,t+1}) = a_1 + a_2\sigma + a_3\sigma_t^2 \tag{64}$$

for some a_1, a_2, a_3, and for u_{t+1} being a vector with constant variance-covariance matrix. Similarly, the covariance between N_V and N_{CF} will be a function of σ_t and σ_t^2. The general intuition for this result is that news about long-run volatility is driven by *all* the shocks in the VAR, not just by the innovation to the volatility equation, and therefore the term N_V will generally have time-varying second moments even when the volatility equation is homoskedastic.

What does this imply? Remember that for $\text{Var}_t(m_{t+1} + r_{t+1})$ to be affine in σ_t^2 we need

$$\text{Var}_t(m_{t+1}+r_{t+1}) = (1-\gamma)^2\text{Var}_t(N_{CF,t+1}) + \omega(1-\gamma)\text{Cov}_t(N_{CF,t+1}, N_{V,t+1}) + \frac{\omega^2}{4}\text{Var}_t(N_{V,t+1}) = f + \omega\sigma_t^2 \tag{65}$$

for some coefficients f and ω (this is analogous to equation (18) with the addition of a constant, f). Under the case considered above, the left-hand side will depend on σ_t in addition to σ_t^2 and a constant. Setting the σ_t term to zero requires additional restrictions on the parameters f and ω and their relation with the news terms; otherwise the proportionality of N_{RISK} and N_V is violated. We consider these restrictions below in greater detail.

Suppose, instead, that by homoskedasticity we mean that N_V itself has constant variance: i.e., $\text{Var}_t(\lambda_v'\eta_{t+1}) = c$, a constant. To obtain this, the vector λ_v must be loading *only* on VAR innovations that are homoskedastic. Even in this case, we can show that a σ_t term will appear on the left-hand side of eq. (65). To see why, note that $N_{CF} = \lambda_{CF}'\eta_{t+1}$, and at least some of the elements of η must depend on σ_t (otherwise, the whole model would be homoskedastic and time-varying volatility would be irrelevant). For simplicity, consider the case $N_{CF} = \sigma_t\lambda_{CF}'u_{t+1}$, which is also the case considered in BKSY. We will have

$$\text{Cov}_t(N_{V,t+1}, N_{CF,t+1}) = \text{Cov}_t(\lambda_{CF}'\sigma_t u_{t+1}, \lambda_V' u_{t+1}) = h\sigma_t \tag{66}$$

for a scalar $h = \text{Cov}(\lambda_{CF}'u_{t+1}, \lambda_V'u_{t+1})$. Eq. (65) then reduces to

$$(1-\gamma)^2\lambda_{CF}'\Sigma\lambda_{CF}\sigma_t^2 + \omega(1-\gamma)h\sigma_t + \frac{\omega^2}{4}c = f + \omega\sigma_t^2 \tag{67}$$

Matching coefficients requires that $\omega(1-\gamma)h = 0$, which can be possible only if either the price of volatility risk is 0 ($\omega = 0$), or if N_{CF} and N_V are uncorrelated. We note that the latter assumption is counterfactual since these news series are negatively correlated in the data.

We conclude that under the assumption of homoskedasticity it will not generally be possible to write $V_t(m_{t+1} + r_{t+1})$ as an affine function of σ_t^2, and therefore generally it will not be the case that N_{RISK} is proportional to N_V.

A similar intuition can be obtained by looking at the conditions for the existence of the value function, in the special case with $\psi = 1$ analyzed above. Suppose that

$$\sigma_{t+1}^2 = s + d\sigma_t^2 + x\epsilon_{t+1} \tag{68}$$

$$\Delta c_{t+1} = k\sigma_t \eta_{t+1} \tag{69}$$

so that σ_t scales the volatility of consumption growth but not its own. Conjecturing that $v_t = a + b\sigma_t^2$ and substituting, we find:

$$a + b\sigma_t^2 = \delta(a + bc + bd\sigma_t^2) + \delta(1-\gamma)0.5[b^2x^2 + k^2\sigma_t^2 + 2bxk\theta\sigma_t]. \tag{70}$$

In the right-hand side we now have a term proportional to σ_t (and not only σ_t^2): $bxk\theta\sigma_t$. For the coefficients on the two sides to match, we need to have:

$$bxk\theta = 0. \tag{71}$$

Therefore, for the value function to have a solution of the form $a + b\sigma_t^2$ we need that either the value function does not depend at all on volatility ($b = 0$), one of the shocks has zero variance ($x = 0$ or $k = 0$), or shocks to volatility and shocks to consumption growth are *uncorrelated* ($\theta = 0$). The latter assumption would imply, counterfactually, that N_V and N_{CF} should be uncorrelated, while they are clearly strongly negatively correlated in the data. Unless one of these conditions is met, the value function cannot be written as an affine function of σ_t^2. And in this case $\mathrm{Var}_t(m_{t+1}+r_{t+1})$, which is proportional to $\mathrm{Var}_t(v_{t+1}+\Delta c_{t+1})$, will not be proportional to σ_t^2, which again implies that the terms N_V and N_{RISK} will not be proportional to each other.

5.2 Simulations of homoskedastic and heteroskedastic volatility processes

In our specification, the conditional variance term σ_t^2 (EVAR) is modeled as one of the elements of the VAR state vector. This means that in samples generated by the VAR σ_t^2 could become negative, an issue that has been previously highlighted in the literature. As in the previous literature, in simulations we replace negative or zero values of σ_t^2 with a small positive number.

In this section we show by simulation that our heteroskedastic VAR specification, in which σ_t^2 scales the entire covariance matrix of the innovations, makes the issue of potentially negative volatility much less severe than the homoskedastic counterpart. This is especially true when the process is sampled at higher frequencies: as we converge towards the continuous time limit, volatility becomes less likely to attain or cross the boundary of zero, and when it does, it crosses the boundary by a small amount.

Starting from our estimated quarterly VAR, we simulate both a homoskedastic version (where the VAR innovations e_{t+1} have a constant covariance matrix equal to the sample unconditional covariance matrix) and the heteroskedastic version we employ in the paper (where the innovations e_{t+1} have a covariance matrix of the form $\sigma_t^2\Sigma$). We then derive the VAR sampled at higher frequencies, and simulate it for 10,000 quarters at quarterly, monthly, and daily frequencies. As we simulate the process, we reset σ_t^2 every time it falls below 0.

For each simulation, we compute the fraction of periods in which σ_t^2 falls below zero, N_0. We also compute the total amount that was added to the process σ_t^2 over the simulated path to ensure it stayed positive everywhere. We refer to this quantity as B_0. To understand the meaning of these quantities, it is useful to consider that in the continuous time limit, a diffusion with reflective boundary at 0 spends zero time at the boundary. The processes we simulate, in which σ_t^2 is reset at **1e-9** when it crosses the boundary, does not converge exactly to a diffusion with reflective boundary in the continuous-time limit, but its behavior is similar to it, so that it would not be surprising if *both* the homoskedastic model and the heteroskedastic model tend to hit the boundary a smaller and smaller fraction of time as we increase the sampling frequency ($N_0 \to 0$). This implies that just counting the amount of times the process needs to be reset (N_0) will not provide a very meaningful diagnostic as we increase the sampling frequency.

Instead, the total amount that needs to be added to the process to ensure its positivity, B_0, converges to a limiting stochastic process, in a reflective diffusion (Karatzas and Shreve 1988). This quantity tells us *by how much* the theoretical process tends to violate the boundary, and therefore is a measure of how well the theoretical VAR (in which this boundary is not taken into account explicitly) can approximate the true process (in which the boundary can never be crossed).

The following table reports the quantities N_0 and B_0 for a 10,000 quarter simulation at different frequencies, and for the two models, homoskedastic and heteroskedastic.

	Quarterly		Monthly		Daily	
	Homosk.	Heterosk.	Homosk.	Heterosk.	Homosk.	Heterosk.
N_0: fraction of times $\sigma_t^2 < 0$	8.8%	3.3%	6.2%	1.7%	1.3%	0.2%
B_0: total amount added to σ_t^2	2.37	0.38	3.65	0.21	4.79	0.02

The Table shows that while for both processes the fraction of times where σ_t^2 needs to be reset, N_0, falls as the frequency increases, the total amount by which σ_t^2 needs to be increased to ensure it stays positive is in fact increasing for the homoskedastic model, while it's decreasing for the heteroskedastic model. As a reference point, the unconditional mean of σ_t^2 is 0.008. So in a daily simulation for 10,000 quarters, the homoskedastic process needs to be increased by 600 times the unconditional mean to stay positive; the heteroskedastic process by only 2.5 times the mean over the entire simulation. This shows that the heteroskedastic process crosses the boundary with low probability and by low amounts as the sampling frequency increases.

5.2.1 Derivations of the higher-frequency VAR representation

We start from the estimated quarterly model:

$$x_{t+1} = a + Gx_t + e_{t+1}. \tag{72}$$

The covariance matrix of e_{t+1} is either constant (and we refer to it just as Σ, empirically the covariance matrix of the unscaled errors) or is stochastic, $\sigma_t^2\Sigma$, where then we refer to Σ as the covariance matrix of the scaled errors.

We start by noting that changing the sampling frequency will not affect the unconditional mean of the model. Therefore, we can rewrite the model in terms of deviations from the mean. The unconditional mean of x_t is:

$$\overline{x} = (I - G)^{-1}a \tag{73}$$

We can then write the model in deviations from the mean as:

$$(x_{t+1} - \overline{x}) = G(x_t - \overline{x}) + e_{t+1}, \tag{74}$$

or:

$$y_{t+1} = Gy_t + e_{t+1} \tag{75}$$

with $y_t = x_{t+1} - \overline{x}$. From now on, we consider the model directly in terms of deviations from the mean.

5.2.2 Scaling the homoskedastic model

Let us now assume that the variance of the e_{t+1} error is Σ, where Σ is the covariance matrix of the unscaled errors. To derive the scaling in the homoskedastic case, suppose we start at a certain frequency and we want to sample less frequently, say every n periods. Then we can write:

$$y_{t+1} = Gy_t + e_{t+1}$$
$$y_{t+2} = G^2 y_t + Ge_{t+1} + e_{t+2}$$
$$\dots$$
$$y_{t+n} = G^n y_t + \sum_{j=0}^{n-1} G^j e_{t+j+1}. \quad (76)$$

The process sampled at frequency $\frac{1}{n}$ has errors $v_{t,t+n} = \sum_{j=0}^{n-1} Ge_{t+j+1}$. The covariance matrix of these errors is:

$$V(v_{t,t+n}) = \sum_{j=0}^{n-1} \left(G^j \Sigma G^{j'} \right). \quad (77)$$

When sampling less frequently, we obtain a VAR with transition matrix G^n and variance of the errors $\sum_{j=0}^{n-1} (G^j \Sigma G^j)$. Conversely, when sampling more frequently, we need a VAR with transition matrix $G^{\frac{1}{n}}$, and a variance-covariance of the errors $\tilde{\Sigma}$ such that:

$$\Sigma = \sum_{j=0}^{n-1} \left(G^{\frac{j}{n}} \tilde{\Sigma} G^{\frac{j}{n}'} \right). \quad (78)$$

$\tilde{\Sigma}$ can be solved by noting that:

$$vec(ABC) = (C' \otimes A) vec(B), \quad (79)$$

so that

$$vec(G^{\frac{j}{n}} \tilde{\Sigma} G^{\frac{j}{n}}) = (G^{\frac{j}{n}} \otimes G^{\frac{j}{n}}) vec(\tilde{\Sigma}). \quad (80)$$

We can rewrite:

$$vec(\Sigma) = \left[\sum_{j=0}^{n-1} (G^{\frac{j}{n}} \otimes G^{\frac{j}{n}}) \right] vec(\tilde{\Sigma}). \quad (81)$$

So we can find

$$vec(\tilde{\Sigma}) = \left[\sum_{j=0}^{n-1} (G^{\frac{j}{n}} \otimes G^{\frac{j}{n}}) \right]^{-1} vec(\Sigma). \quad (82)$$

There are two practical considerations to take into account. First, $G^{\frac{1}{n}}$ may not be real. In that case, we take the real part of it. Second, $\tilde{\Sigma}$ is not guaranteed to be positive semidefinite, i.e. a valid covariance matrix. In that case, we replace each negative eigenvalues with a small positive number (1e-6). We have verified by simulation that both of these modifications have small effects on the results.

5.2.3 Scaling the heteroskedastic model

The heteroskedastic model has one additional difficulty: that in theory the volatility is changing within each period, continuously. Ignoring this, we can simply find the time-scaled variance of the scaled shocks by solving:

$$vec(\tilde{\Sigma}) = \left[\sum_{j=0}^{n-1} (G^{\frac{j}{n}} \otimes G^{\frac{j}{n}}) \right]^{-1} vec(\Sigma), \tag{83}$$

and then construct the conditional variance as $\sigma_t \tilde{\Sigma}$.

5.3 Comparison of the empirical results

In Appendix Table 6 we estimate different versions of the VAR that include our baseline estimate, our replication of BKSY's VAR, and various combination of the two specifications. For each specification, the Table reports the properties of the news terms. The main differences between BKSY's and our VAR are: 1) BKSY estimate the VAR at the yearly frequency, we estimate it at the quarterly frequency; 2) BKSY use monthly returns to construct RVAR, we use daily returns; 3) the state variable used by BKSY are partly different than the ones we use: in addition to RVAR, BKSY use log dividend growth, log price-dividend ratio, term spread, long-term interest rate, and default spread; we use the market return, the log PE ratio, the value spread, the default spread, and the t-bill rate.

6 Construction of the Test Portfolios

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.[5]

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

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

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.

We also 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.

$$\sum_{j=1}^{3} r_{i,t+j} = b_0 + b_{r_M} \sum_{j=1}^{3} r_{M,t+j} + b_{\Delta VAR} \sum_{j=1}^{3} \Delta VAR_{t+j} + \varepsilon_{i,t+3} \tag{84}$$

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.

Finally, we 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.

Appendix Table 7 reports the average excess returns on our test assets. Appendix Table 8 reports the ICAPM betas for the size and book-to-market sorted portfolios. Appendix Table 9 reports the ICAPM betas for the characteristic and risk-sorted portfolios.

7 Changing Volatility Beta of the Aggregate Stock Market

In the paper we find that the average β_V of the 25 size- and book-to-market portfolios changes sign from the early to the modern subperiod. Over the 1931-1963 period, the average β_V is -0.10 while over the 1964-2011 period this average becomes 0.06. Of course, given the strong positive link between PE and volatility news documented in the paper, one should not be surprised that the market's β_V can be positive. Moreover, in Appendix Table 1 we show that the correlation between PE and some of the key variables driving $EVAR$ changes from one subperiod to the other. Nevertheless, we study this change in sign more carefully.

Appendix Figure 4 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 the paper, the relation between the market return and N_V is negative in the early period and positive in the modern period.[6] 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

[6] 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.

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, Appendix Figure 4 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.

7.1 Model estimation

We now turn to pricing the cross section of excess returns on our test assets. We estimate our model's single parameter via GMM, using the moment condition (44). For ease of exposition, we report our results in terms of the expected return-beta representation from equation (**??**), rescaled by the variance of market return innovations as in section 5.2:

$$\overline{R}_i - \overline{R}_j = g_1 \widehat{\beta}_{i,CF_M} + g_2 \widehat{\beta}_{i,DR_M} + g_3 \widehat{\beta}_{i,V_M} + e_i, \tag{85}$$

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

We evaluate the performance of five asset pricing models, all estimated via GMM: 1) the traditional CAPM that restricts cash-flow and discount-rate betas to have the same price of risk and sets the price of variance risk to zero; 2) the two-beta intertemporal asset pricing model of CV (2004) that restricts the price of discount-rate risk to equal the variance of the market return and again sets the price of variance risk to zero; 3) our three-beta intertemporal asset pricing model that restricts the price of discount-rate risk to equal the variance of the market return and constrains the prices of cash-flow and variance risk to be related by equation (**??**), 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.

7.1.1 Model estimates with characteristic-sorted portfolios

Table 10 reports separate results for the early sample period 1931-1963 (Panel A) and the modern sample period 1963-2011 (Panel B), using 25 size- and book-to-market-sorted portfolios and the T-bill rate as test assets. The table has five columns, one for each of our asset pricing models. The first six rows of each panel in Table 10 are divided into three sets of two rows. The first set of two rows corresponds to the premium on cash-flow beta, the second set to the premium on discount-rate beta, and the third set to the premium on volatility beta. Within each set, the first row reports the point estimate in fractions per quarter, and the second row reports the corresponding standard error. Below the premia estimates, we report the R^2 statistic for a cross-sectional regression of average market-adjusted returns on our test assets onto the fitted values from the model as well as the J statistic. In the final two rows of each panel, we report the implied risk-aversion coefficient, γ, which can be recovered as g_1/g_2, as well as the sensitivity of news about risk to news about market variance, ω, which can be recovered as $-2g_3/g_2$.

Table 10 Panel A shows that in the 1931-1963 period, all our models explain the cross section of stock returns reasonably well. The cross-sectional R^2 statistic is 64% for the CAPM, 66% for the two-beta ICAPM, and 67% for our three-beta ICAPM. Consistent with the claim that the three-beta model does a good job describing the cross section, the constrained and the unrestricted factor model barely improve pricing relative to the three-beta ICAPM in Panel A. Despite this apparent success, all models are rejected based on the standard J test. This may not be surprising, given that even the empirical three-factor model of Fama and French (1993) is rejected by this test.

Results are very different in the 1963-2011 period. Table 10 Panel B shows that in this period, the CAPM does a very poor job of explaining cross-sectional variation in average market-adjusted returns on size and value portfolios: its cross-sectional R^2 is strongly negative at -50%. The two-beta CV (2004) model does a much better job describing the cross section of average returns than the CAPM, with a cross-sectional R^2 of 45%. However, the implied coefficient of risk aversion is arguably extreme at 16.5, and much larger than the value of 6.4 estimated in the early subperiod.

The three-beta model explains slightly more cross-sectional variation than the two-beta model, delivering an R^2 of 48%. Importantly, the estimated coefficient of relative risk aversion is estimated at 7.2, a moderate value that is reasonably similar to the estimate of 5.2 from the early subperiod. The value of ω that corresponds to this estimate of risk aversion is 24.9. As before, all models are rejected based on the J statistic.

The modest size of the increase in R^2 delivered by the three-beta ICAPM is because of the derived link between γ and ω. We illustrate this fact by considering in Panel B a partially-constrained factor model that removes the constraint linking γ and ω but retains the constraint on the discount-rate beta premium. The cross-sectional R^2 for this model increases from 48% to 76%, and the risk prices for γ and ω remain economically large and of the right sign. The γ implied by the partially-constrained model is 15.1, and the implied ω is 27.3. Thus, compared to our fully-constrained model, the data prefer a higher γ rather than a higher ω .

8 Implications for Consumption Growth

In this section we derive the model implications for consumption growth and show how to represent the stochastic discount factor in terms of consumption growth, news about future consumption growth, and consumption volatility.

8.1 Implied consumption growth in the model and in the data

8.1.1 Analytical results

Following Campbell (1993), in this paper we substitute consumption out of the pricing equations using the intertemporal budget constraint. However the model does have interesting implications for the implied consumption process. From equation (4) in the text and the identity $r_{t+1} - \mathrm{E}_t r_{t+1} = (\Delta c_{t+1} - \mathrm{E}_t \Delta c_{t+1}) + (h_{t+1} - \mathrm{E}_t h_{t+1})$, we can derive the expression:

$$\Delta c_{t+1} - E_t \Delta c_{t+1} = (r_{t+1} - E_t r_{t+1}) - (\psi - 1) N_{DR,t+1} - (\psi - 1) \frac{1}{2} \frac{\omega}{1-\gamma} N_{V,t+1}. \qquad (86)$$

The first two components of the equation for consumption growth are the same as in the homoskedastic case. An unexpectedly high return of the wealth portfolio has a one-for-one effect on consumption. An increase in expected future returns increases today's consumption if $\psi < 1$, as the low elasticity of intertemporal substitution induces the representative investor to consume today (the income effect dominates). If $\psi > 1$, instead, the same increase induces the agent to reduce consumption to better exploit the improved investment opportunities (the substitution effect dominates).

The introduction of time-varying conditional volatility adds an additional term to the equation describing consumption growth. News about high future risk is news about a deterioration of future investment opportunities, which is bad news for a risk-averse investor ($\gamma > 1$). When $\psi < 1$, the representative agent will reduce consumption and save to ensure adequate future consumption. An investor with high elasticity of intertemporal substitution, on the other hand, will increase current consumption and reduce the amount of wealth exposed to the future (worse) investment opportunities.

Using estimates of the news terms from our VAR model, we can explore the implications of the model for consumption growth. As shown in the text, the three shocks that drive innovations in consumption growth ($r_{t+1} - E_t r_{t+1}$, $N_{DR,t+1}$, $N_{V,t+1}$) can all be expressed as functions of the vector of innovations $\sigma_t u_{t+1}$. The conditional variance of consumption growth, $\text{Var}_t(\Delta c_{t+1})$, will then be proportional to the conditional variance of returns, $\text{Var}_t(r_{t+1})$; similarly, the conditional standard deviation of consumption growth will be proportional to the conditional standard deviation of returns. As a consequence, the ratio of the standard deviations,

$$A(\gamma, \psi) \equiv \frac{\sqrt{\text{Var}_t(\Delta c_{t+1})}}{\sqrt{\text{Var}_t(r_{t+1})}} \tag{87}$$

will be a constant that depends on the model parameters γ and ψ as well as on the unconditional variances and covariances of the innovation vector u_{t+1}, which we obtain by estimating the VAR.

Appendix Figure 5 plots the coefficient $A(\gamma, \psi)$ for different values of γ and ψ for the homoskedastic case (left panel), and for the heteroskedastic case (right panel). In each panel, we plot $A(\gamma, \psi)$ as γ varies between 0 and the maximum possible value of γ, for different values of ψ. Each line corresponds to a different ψ between 0.5 and 1.5; when $\psi = 1$ the value of $A(\gamma, \psi)$ is always equal to 1 since in that case the volatility of consumption growth is equal to the volatility of returns.

As expected, in the homoskedastic case (left panel), the variance of consumption growth does not depend on γ but only on ψ. It is rising in ψ because our VAR estimates imply that the return on wealth is negatively correlated with news about future expected returns $N_{DR,t+1}$, that is, wealth returns are mean-reverting. This confirms results reported in Campbell (1996). Once we add stochastic volatility (right panel), as γ increases the volatility of consumption growth increases for all values of ψ as long as $\psi > 1$, while for values of $\psi < 1$, the effect depends on ψ. The reason for this is that the variance of consumption growth depends on the variances and covariances of the three terms that add up to consumption growth as shown in equation (86). Whether $A(\gamma, \psi)$ increases with γ or not depends on the relative magnitude of these variances and covariances, which in turn depends on ψ.

Overall, Appendix Figure 5 shows that including stochastic volatility makes little difference to the variance of consumption growth for the range of γ in which the model admits a solution. And γ has a relatively minor effect on the variance of consumption growth, which continues to depend primarily on ψ.

8.1.2 Implied and measured aggregate consumption and cash flows

Next we compare the implied consumption innovations ($\Delta c_{t+1} - E_t \Delta c_{t+1}$) to observed innovations in real log aggregate consumption growth, as well as stockholders' and top stockholder's log consumption growth. We construct aggregate consumption growth using nondurable and services data from the BEA, and obtain stockholders' and top stockholders' data from Malloy, Moskowitz, and Vissing-Jørgensen (2009). We construct consumption innovations by taking the residuals of an AR(1) regression for each series. We work here with yearly data, from 1930 to 2011 for all series except for stockholders' consumption, which is available only between 1982 and 2004.

Panel A of Appendix Table 11 reports the standard deviations of innovations in aggregate consumption in the full sample and in the subsample 1982-2004 for which stockholders' consumption data is available; stockholders' and top stockholders' consumption (only for the period 1982-2004); and implied innovations in consumption under the three calibrations for ψ (0.5, 1, 1.5). The table shows that the implied consumption innovation series are more volatile than aggregate consumption innovations, but roughly of the same order of magnitude as top stockholders' consumption.

Panel B of Appendix Table 11 reports the correlations between the consumption innovation series implied by our model and the observed consumption series (for aggregate consumption, we compute the correlations both in the full sample and in the subsample 1982-2004 for which we observe stockholders' consumption as well). The table shows that our implied consumption series are positively correlated with the realizations of aggregate consumption in the full sample (with a correlation of about 0.3). While the correlation of aggregate consumption with implied consumption is weaker in the subsample 1982-2004, top stockholders' consumption (only available in this time period) correlates more strongly than aggregate consumption with the implied consumption innovations. This is consistent with the interpretation that the investor whose first-order conditions we study in this paper is a long-term investor fully invested in the market portfolio, whose consumption lines up more strongly with the stockholders' consumption series.

Appendix Figure 6 shows the time series of aggregate and stockholder's consumption growth, as well as the time series of implied consumption growth (with $\psi = 0.5$). Both the higher volatility of stockholder's consumption relative to aggregate consumption and the higher correlation with implied consumption are clearly visible in the figure.

8.2 Consumption-based representation of the SDF

As shown in the previous section, we can express consumption growth as:

$$\Delta c_{t+1} - E_t \Delta c_{t+1} = (r_{t+1} - E_t r_{t+1}) - (\psi - 1) N_{DR,t+1} - (\psi - 1) \frac{1}{2} \frac{\omega}{1-\gamma} N_{V,t+1}. \tag{88}$$

Note that we also have:

$$m_{t+1} - E_t m_{t+1} = -\gamma N_{CF,t+1} + N_{DR,t+1} + \frac{1}{2} \omega N_{V,t+1}. \tag{89}$$

We can rewrite the consumption equation as:

$$\Delta c_{t+1} - E_t \Delta c_{t+1} = N_{CF,t+1} - \psi N_{DR,t+1} - (\psi - 1)\frac{1}{2}\frac{\omega}{1-\gamma} N_{V,t+1} \tag{90}$$

and, rearranging:

$$\psi N_{DR,t+1} = -(\Delta c_{t+1} - E_t \Delta c_{t+1}) + N_{CF,t+1} - (\psi - 1)\frac{1}{2}\frac{\omega}{1-\gamma} N_{V,t+1}. \tag{91}$$

$$N_{DR,t+1} = -\frac{1}{\psi}(\Delta c_{t+1} - E_t \Delta c_{t+1}) + \frac{1}{\psi} N_{CF,t+1} - \frac{1}{\psi}(\psi - 1)\frac{1}{2}\frac{\omega}{1-\gamma} N_{V,t+1}. \tag{92}$$

Substitute into the SDF:

$$m_{t+1} - E_t m_{t+1} = -\gamma N_{CF,t+1} + \left\{ -\frac{1}{\psi}(\Delta c_{t+1} - E_t \Delta c_{t+1}) + \frac{1}{\psi} N_{CF,t+1} - \frac{1}{\psi}(\psi - 1)\frac{1}{2}\frac{\omega}{1-\gamma} N_{V,t+1} \right\} + \frac{1}{2}\omega N_{V,t+1} \tag{93}$$

or:

$$m_{t+1} - E_t m_{t+1} = -\frac{1}{\psi}(\Delta c_{t+1} - E_t \Delta c_{t+1}) - (\gamma - \frac{1}{\psi}) N_{CF,t+1} + \frac{1}{2}\omega \left\{ -\frac{1}{\psi}(\psi - 1)\frac{1}{1-\gamma} + 1 \right\} N_{V,t+1} \tag{94}$$

Note that:

$$\frac{\psi - 1}{1-\gamma} = \frac{\psi}{\left(\frac{1-\gamma}{1-\frac{1}{\psi}}\right)} \tag{95}$$

so we have:

$$\left\{ -\frac{1}{\psi}(\psi - 1)\frac{1}{1-\gamma} + 1 \right\} = \left\{ -\frac{1}{\psi}\frac{\psi}{\left(\frac{1-\gamma}{1-\frac{1}{\psi}}\right)} + 1 \right\} = \left\{ -\frac{1}{\left(\frac{1-\gamma}{1-\frac{1}{\psi}}\right)} + 1 \right\} = \left\{ -\frac{1-\frac{1}{\psi}}{1-\gamma} + 1 \right\} = -\frac{1}{\theta} + 1$$

$$= \frac{\theta - 1}{\theta} \tag{96}$$

So to conclude we have:

$$m_{t+1} - E_t m_{t+1} = -\frac{1}{\psi}(\Delta c_{t+1} - E_t \Delta c_{t+1}) - (\gamma - \frac{1}{\psi}) N_{CF,t+1} + \frac{1}{2}\omega \left(\frac{\theta - 1}{\theta} \right) N_{V,t+1}. \tag{97}$$

As expected, when $\theta = 1$, the price of N_V is 0. The N_V reported here is the volatility of news about the volatility of *returns*. Next, we express the SDF in terms of news about the volatility of consumption growth N_{CV}.

Recall that:

$$\omega\sigma_t^2 = V_t((1-\gamma)N_{CF,t+1} + \frac{1}{2}\omega N_{V,t+1}), \tag{98}$$

$$\sigma_t^2 = V_t(r_{t+1} - E_t r_{t+1}). \tag{99}$$

We have:

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

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

$$= (r_{t+1} - E_t r_{t+1}) + (\psi-1)(r_{t+1} - E_t r_{t+1}) - (\psi-1)N_{CF,t+1} - (\psi-1)\frac{1}{2}\frac{\omega}{1-\gamma}N_{V,t+1}$$

$$= \psi(r_{t+1} - E_t r_{t+1}) - \frac{(\psi-1)}{1-\gamma}\left[(1-\gamma)N_{CF,t+1} + \frac{1}{2}\omega N_{V,t+1}\right]. \tag{100}$$

So:

$$V_t(\Delta c_{t+1} - E_t\Delta c_{t+1}) = \psi^2\sigma_t^2 + \left(\frac{(\psi-1)}{1-\gamma}\right)^2\omega\sigma_t^2 +$$

$$-\psi\frac{(\psi-1)}{1-\gamma}Cov_t(r_{t+1} - E_t r_{t+1}, (1-\gamma)N_{CF,t+1} + \frac{1}{2}\omega N_{V,t+1}). \tag{101}$$

Note that while ω only depends on the covariance between N_{CF} and N_V, this conditional variance depends on the covariance between all three news terms N_{DR}, N_{CF} and N_V. If we call this coefficient k, we have:

$$V_t(\Delta c_{t+1} - E_t\Delta c_{t+1}) = k\sigma_t^2, \tag{102}$$

and

$$N_{CV} = kN_V \tag{103}$$

then we can rewrite:

$$m_{t+1} - E_t m_{t+1} = -\frac{1}{\psi}(\Delta c_{t+1} - E_t\Delta c_{t+1}) - (\gamma - \frac{1}{\psi})N_{CF,t+1} + \frac{1}{2}\omega\left(\frac{\theta-1}{\theta}\right)\frac{1}{k}N_{CV,t+1} \tag{104}$$

or, more simply, we can call $\eta = \frac{\omega}{k}$ and obtain the formula reported in the text.

9 Implications for the Risk-Free Rate

In this section we report the implied risk-free rate r_f in the model as well as N_{rf} (news about the future risk-free rate), and plot their time series using our model estimates.

9.1 Time series of r_f and N_{rf}

Starting from the equation for the conditional risk premium of the market:

$$E_t r^M_{t+1} + \frac{1}{2}\sigma^2_{Mt} - r^f_{t+1} = \gamma Cov_t(r^M_{t+1}, N_{CF,t+1}) - Cov_t(r^M_{t+1}, N_{DR,t+1}) - \frac{1}{2}\omega Cov_t(r^M_{t+1}, N_{V,t+1}) \tag{105}$$

We derive the log risk-free rate r^f_{t+1} as:

$$r^f_{t+1} = E_t r^M_{t+1} - H\sigma^2_t, \tag{106}$$

where

$$H = \gamma e_1' \Sigma x_{CF} - e_1' \Sigma x_{DR} - \frac{1}{2}\omega e_1' \Sigma x_{NV} - \frac{1}{2} \tag{107}$$

includes both the risk premium and the volatility adjustment due to the Jensen adjustment for log returns.

Here x_{CF}, x_{DR} and x_{NV} are the loadings of the three news terms on the vector of innovations, and Σ is the scaled covariance matrix of innovations. In our data, and using our estimates of γ and ω, we find $H = 2.27$.

Appendix Figure 7 plots $E_t r^M_{t+1}$ and r^f_{t+1}. The difference between the two lines is, of course, $H\sigma^2_t$. The risk-free rate implied by the model is relatively volatile (standard deviation of 2.4% per quarter). The graph clearly shows periods where r_f and $E_t r^M_{t+1}$ are close to each other (times of low market volatility), like in the 1950s and 1960s; but also periods when the two diverge significantly, as in the Great Depression and the Great Recession, where volatility is higher. The implied risk-free rate turns sharply negative in a few occasions – the Great Depression, during the tech boom in the 1990s-2000s (prices were high and expected returns were low) and at the onset of the financial crisis, when volatility spiked.

Risk-free news is:

$$N_{rf,t+1} = N_{DR,t+1} - H \cdot N_{V,t+1} \tag{108}$$

Appendix Figure 8 shows the time-series of N_{rf} (normalized to have standard deviation of 1 and then smoothed as in Figure 2 in the text). Risk-free rate news is correlated with changes in the risk-free rate, but, remarkably, does not show a large drop at the beginning of the financial crisis. This is because the news reflects all the future dynamics of the risk-free rate, so that a temporary spike in the risk-free rate has a smaller effect on N_{rf} than on r_f itself.

Finally, Appendix Table 12 reports the correlations among all these news terms (with standard deviations on the diagonal). Not surprisingly, risk-free rate news is positively correlated with discount rate news, and negatively correlated with news about future variance.

9.2 Derivations

Recall that:

$$E_t r^M_{t+1} + \frac{1}{2}\sigma^2_{Mt} - r^f_{t+1} = \gamma Cov_t(r^M_{t+1}, N_{CF,t+1}) - Cov_t(r^M_{t+1}, N_{DR,t+1}) - \frac{1}{2}\omega Cov(r^M_{t+1}, N_{V,t+1}) \tag{109}$$

Now, call x_{CF}, x_{DR} and x_{NV} the loadings of the news terms on the vector u_t; e_1 is a vector of zeroes with 1 as the first element, so that:

$$Cov_t(r^M_{t+1}, N_{CF,t+1}) = e_1' \Sigma x_{CF} \sigma^2_t \tag{110}$$

where Σ is the covariance matrix of scaled news terms. Then we can write:

So we can write:

$$\begin{aligned} r^f_{t+1} &= E_t r_{t+1} + \frac{1}{2}\sigma^2_t - \gamma e_1'\Sigma x_{CF}\sigma^2_t + e_1'\Sigma x_{DR}\sigma^2_t + \frac{1}{2}\omega e_1'\Sigma x_{NV}\sigma^2_t \\ &= E_t r_{t+1} + (\frac{1}{2} - \gamma e_1'\Sigma x_{CF} + e_1'\Sigma x_{DR} + \frac{1}{2}\omega e_1'\Sigma x_{NV})\sigma^2_t. \end{aligned} \tag{111}$$

We define:

$$N_{rf,t+1} = (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j r^f_{t+j+1}$$

$$= (E_{t+1} - E_t)\sum_{j=1}^{\infty} \rho^j \left[E_{t+j} r_{t+j+1} + (\frac{1}{2} - \gamma e_1'\Sigma x_{CF} + e_1'\Sigma x_{DR} + \frac{1}{2}\omega e_1'\Sigma x_{NV})\sigma^2_{t+j} \right], \tag{112}$$

so

$$N_{rf,t+1} = N_{DR,t+1} + (\frac{1}{2} - \gamma e_1'\Sigma x_{CF} + e_1'\Sigma x_{DR} + \frac{1}{2}\omega e_1'\Sigma x_{NV}) N_{V,t+1}. \tag{113}$$

10 Robustness

Appendix Table 13 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 is 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 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 Appendix

Table 3 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 an OLS VAR, 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 first major finding (a negative β_V for HML) is generally robust to using different methods. However, the use of WLS is critical for successful ICAPM pricing.

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, ICAPM R^2s remain high in the modern period for quarterly VARs.

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 Appendix Table 3 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.[7] In unreported results, we find 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, R^{Tbill}, 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

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

that our finding of a negative β_V for HML and successful 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 strongly correlated with their in-sample counterparts. Specifically, the correlations are 0.36, 0.43, and 0.63 for the cash-flow, discount-rate, and volatility news terms respectively over the 1963:3-2011:4 subperiod. Appendix Table 13 Panel E documents that the pricing of the out-of-sample news terms is very consistent with the full-sample results. In fact, the out-of-sample test of the ICAPM has a much higher R^2. We next allow for a structural break between the early and modern periods in the coefficients of the return and volatility regressions in the VAR. Specifically, we interact all coefficients in these regressions in both the first and second stage with a dummy variable that is one if the observation is in the post-1962 subsample. Thus, we effectively split the sample for these two key regressions of the VAR. We continue to focus on the modern period. We find that the news terms from these sample splits are highly correlated with their baseline counterparts. We again find that HML's β_V is negative. As with our baseline specification, the modern period cross-sectional R^2 is approximately 48%.

In unreported results, we have also experimented with allowing all 42 coefficients in the VAR's six regressions to vary across the early and modern periods. Care has to be taken here as more persistent variables rightly receive more weight in the construction of the news terms yet estimates of persistence are known to be biased downwards in finite samples with the size of the bias proportional to the inverse of the sample size (Kendall 1954). Nevertheless, our main finding, that HML's β_V is negative, continues to hold in these specifications. However, in the modern subsample, we estimate a less persistent process for VS, the key variable in Campbell and Vuolteenaho's (2004) finding that HML's β_{CF} is positive and, as a consequence, find that HML's β_{CF} is negative in that specification.

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 two specific delevered portfolios we examine have 80% or 60% 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 declining slowly. 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 10.2 while delevering by 40% requires risk aversion of 14.5.

Panels G and H 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 G 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 G 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 G 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 G 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.

Panel H 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 22% of HML's returns in the modern period. The three news factors together explain slightly over 32%. 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).

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.

Bansal, Ravi, Dana Kiku, Ivan Shaliastovich and Amir Yaron, 2014, "Volatility, the Macroeconomy and Asset Prices", *Journal of Finance* 69:2471–2511.

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.

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.

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

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. 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. 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, 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.

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.

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.

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. 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.

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

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.

Kendall, M.G., 1954, "Note on Bias in the Estimation of Autocorrelation", *Biometrika* 41:403–404.

Lettau, Martin and Sydney C. Ludvigson, 2001, "Consumption, Aggregate Wealth, and Expected Stock Returns", *Journal of Finance* 56, 815–849.

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, Stijn Van Nieuwerburgh, and Adrien Verdelhan, 2013, "The Wealth-Consumption Ratio", *Review of Asset Pricing Studies* 3, 38–94.

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.

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

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.

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.

Table 1: Summary Statistics

This Appendix Table reports descriptive statistics for quarterly observations of the state variables included in the VAR. r_M is the log real return on the CRSP value-weight index. $RVAR$ is the realized variance of within-quarter daily 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. r^{Tbill} is the log three-month Treasury Bill yield. DEF is the default yield spread in percentage points, measured as the difference between the log yield on Moody's BAA bonds and the log yield on Moody's AAA bonds. VS is the small-stock value-spread, the difference in the log book-to-market ratios of small value and small growth stocks. The small-value and small-growth portfolios are two of the six elementary portfolios constructed by Davis et al. (2000). The paper reports the WLS parameter estimates of a constrained regression forecasting $RVAR$ with lagged values of these state variables; the forecasted values from that regression are the state variable $EVAR$ used in the second stage of the estimation and described below. Panel A reports descriptive statistics of these state variables over the full sample period 1926.2-2011.4, 343 quarterly data points. Panel B reports descriptive statistics of these state variables over the early sample period 1926.2-1963.2, 149 quarterly data points. Panel C reports descriptive statistics of these state variables over the modern sample period 1963.3-2011.4, 194 quarterly data points. "Stdev." denotes standard deviation and "Autocorr." the first-order autocorrelation of the series.

Panel A: Full-Sample Summary Statistics

Variable	Mean	Median	Stdev.	Min	Max	Autocorr.	
r_M	0.016	0.027	0.107	-0.406	0.635	-0.038	
$RVAR$	0.007	0.003	0.012	0.000	0.113	0.524	
$EVAR$	0.007	0.005	0.007	0.000	0.062	0.754	
PE	2.924	2.919	0.379	1.508	3.910	0.965	
r_{Tbill}	0.016	0.014	0.013	0.000	0.063	0.965	
DEF	1.072	0.852	0.671	0.324	5.167	0.901	
VS	1.640	1.517	0.356	1.183	2.685	0.969	
Correlations	$r_{M,t+1}$	$RVAR_{t+1}$	$EVAR_{t+1}$	PE_{t+1}	$r_{Tbill,t+1}$	DEF_{t+1}	VS_{t+1}
$r_{M,t+1}$	1	-0.306	-0.306	0.083	-0.040	-0.138	-0.037
$RVAR_{t+1}$	-0.306	1	0.922	-0.214	-0.189	0.600	0.336
$EVAR_{t+1}$	-0.306	0.922	1	-0.166	-0.229	0.777	0.505
PE_{t+1}	0.083	-0.214	-0.166	1	0.106	-0.597	-0.356
$r_{Tbill,t+1}$	-0.040	-0.189	-0.229	0.106	1	-0.133	-0.482
DEF_{t+1}	-0.138	0.600	0.777	-0.597	-0.133	1	0.645
VS_{t+1}	-0.037	0.336	0.505	-0.356	-0.482	0.645	1
$r_{M,t}$	-0.038	-0.157	-0.156	0.095	-0.014	-0.163	-0.023
$RVAR_t$	0.031	0.524	0.594	-0.215	-0.205	0.568	0.355
$EVAR_t$	0.018	0.609	0.754	-0.169	-0.246	0.717	0.515
PE_t	-0.156	-0.113	-0.067	0.965	0.113	-0.541	-0.350
$r_{Tbill,t}$	-0.031	-0.156	-0.191	0.101	0.965	-0.102	-0.477
DEF_t	0.077	0.520	0.661	-0.582	-0.150	0.901	0.640
VS_t	-0.032	0.338	0.491	-0.358	-0.490	0.619	0.969
Covariances	r_M	$RVAR$	$EVAR$	PE	r_{Tbill}	DEF	VS
r_M	0.0115	-0.0004	-0.0002	0.0032	-0.0001	-0.0100	-0.0015
$RVAR$	-0.0004	0.0001	0.0001	-0.0009	0.0000	0.0047	0.0014
$EVAR$	-0.0002	0.0001	0.0001	-0.0005	0.0000	0.0038	0.0013
PE	0.0032	-0.0009	-0.0005	0.1434	0.0005	-0.1512	-0.0478
r_{Tbill}	-0.0001	0.0000	0.0000	0.0005	0.0002	-0.0012	-0.0022
DEF	-0.0100	0.0047	0.0038	-0.1512	-0.0012	0.4496	0.1538
VS	-0.0015	0.0014	0.0013	-0.0478	-0.0022	0.1538	0.1264

Panel B: 1926-1963 Summary Statistics

Variable	Mean	Median	Stdev.	Min	Max	Autocorr.	
r_M	0.020	0.029	0.128	-0.406	0.635	-0.105	
$RVAR$	0.008	0.003	0.013	0.001	0.091	0.568	
$EVAR$	0.008	0.004	0.009	0.000	0.043	0.812	
PE	2.715	2.723	0.300	1.508	3.502	0.914	
RF	0.006	0.005	0.006	0.000	0.021	0.937	
DEF	1.214	0.820	0.879	0.435	5.167	0.910	
VS	1.838	1.730	0.441	1.236	2.685	0.981	
Correlations	$r_{M,t+1}$	$RVAR_{t+1}$	$EVAR_{t+1}$	PE_{t+1}	$r_{Tbill,t+1}$	DEF_{t+1}	VS_{t+1}
$r_{M,t+1}$	1	-0.233	-0.311	0.128	0.013	-0.217	-0.100
$RVAR_{t+1}$	-0.233	1	0.918	-0.458	-0.123	0.684	0.411
$EVAR_{t+1}$	-0.311	0.918	1	-0.533	-0.185	0.894	0.644
PE_{t+1}	0.128	-0.458	-0.533	1	0.605	-0.727	-0.501
$r_{Tbill,t+1}$	0.013	-0.123	-0.185	0.605	1	-0.340	-0.549
DEF_{t+1}	-0.217	0.684	0.894	-0.727	-0.340	1	0.777
VS_{t+1}	-0.100	0.411	0.644	-0.501	-0.549	0.777	1
$r_{M,t}$	-0.105	-0.107	-0.134	0.131	0.010	-0.170	-0.061
$RVAR_t$	0.046	0.568	0.642	-0.447	-0.170	0.652	0.428
$EVAR_t$	0.025	0.684	0.812	-0.530	-0.237	0.838	0.651
PE_t	-0.239	-0.332	-0.383	0.914	0.605	-0.615	-0.480
$r_{Tbill,t}$	0.001	-0.053	-0.133	0.580	0.937	-0.316	-0.528
DEF_t	0.068	0.667	0.812	-0.704	-0.383	0.910	0.771
VS_t	-0.039	0.410	0.623	-0.494	-0.567	0.749	0.981
Covariances	r_M	$RVAR$	$EVAR$	PE	r_{Tbill}	DEF	VS
r_M	0.0163	-0.0004	-0.0004	0.0048	0.0000	-0.0243	-0.0058
$RVAR$	-0.0004	0.0002	0.0001	-0.0018	0.0000	0.0078	0.0024
$EVAR$	-0.0004	0.0001	0.0001	-0.0014	0.0000	0.0069	0.0025
PE	0.0048	-0.0018	-0.0014	0.0898	0.0010	-0.1911	-0.0657
r_{Tbill}	0.0000	0.0000	0.0000	0.0010	0.0000	-0.0017	-0.0014
DEF	-0.0243	0.0078	0.0069	-0.1911	-0.0017	0.7723	0.3009
VS	-0.0058	0.0024	0.0025	-0.0657	-0.0014	0.3009	0.1945

Panel C: 1963-2011 Summary Statistics

Variable	Mean	Median	Stdev.	Min	Max	Autocorr.	
$RVAR$	0.006	0.004	0.011	0.000	0.113	0.464	
$EVAR$	0.007	0.006	0.006	0.000	0.062	0.653	
PE	3.085	3.114	0.354	2.331	3.910	0.976	
r_{Tbill}	0.023	0.022	0.013	0.000	0.063	0.943	
DEF	0.963	0.855	0.421	0.324	3.167	0.854	
VS	1.488	1.484	0.147	1.183	2.045	0.809	
Correlations	$r_{M,t+1}$	$RVAR_{t+1}$	$EVAR_{t+1}$	PE_{t+1}	$r_{Tbill,t+1}$	DEF_{t+1}	VS_{t+1}
$r_{M,t+1}$	1	-0.414	-0.301	0.102	-0.053	0.008	0.060
$RVAR_{t+1}$	-0.414	1	0.936	-0.005	-0.235	0.480	0.256
$EVAR_{t+1}$	-0.301	0.936	1	0.173	-0.367	0.554	0.377
PE_{t+1}	0.102	-0.005	0.173	1	-0.558	-0.543	0.429
$r_{Tbill,t+1}$	-0.053	-0.235	-0.367	-0.558	1	0.193	-0.213
DEF_{t+1}	0.008	0.480	0.554	-0.543	0.193	1	0.060
VS_{t+1}	0.060	0.256	0.377	0.429	-0.213	0.060	1
$r_{M,t}$	0.064	-0.233	-0.196	0.131	0.010	-0.185	-0.007
$RVAR_t$	0.005	0.464	0.529	-0.011	-0.242	0.441	0.287
$EVAR_t$	0.002	0.507	0.653	0.167	-0.373	0.473	0.391
PE_t	-0.104	0.106	0.257	0.976	-0.548	-0.532	0.421
$r_{Tbill,t}$	-0.028	-0.199	-0.309	-0.567	0.943	0.270	-0.201
DEF_t	0.088	0.265	0.346	-0.524	0.184	0.854	0.044
VS_t	-0.114	0.269	0.370	0.407	-0.231	0.015	0.809
Covariances	r_M	$RVAR$	$EVAR$	PE	r_{Tbill}	DEF	VS
r_M	0.0079	-0.0004	-0.0002	0.0032	-0.0001	0.0002	0.0008
$RVAR$	-0.0004	0.0001	0.0001	0.0000	0.0000	0.0021	0.0004
$EVAR$	-0.0002	0.0001	0.0000	0.0004	0.0000	0.0014	0.0003
PE	0.0032	0.0000	0.0004	0.1255	-0.0025	-0.0807	0.0222
r_{Tbill}	-0.0001	0.0000	0.0000	-0.0025	0.0002	0.0010	-0.0004
DEF	0.0002	0.0021	0.0014	-0.0807	0.0010	0.1768	0.0033
VS	0.0008	0.0004	0.0003	0.0222	-0.0004	0.0033	0.0215

Table 2: VAR Estimation

This Appendix table reports the correlation ("Corr/std") and autocorrelation ("Autocorr.") matrices of both the unscaled and scaled shocks from the second-stage VAR estimated in ICSV Table 1. The sample period for the dependent variables is 1926.3-2011.4, 342 quarterly data points.

Autocorrelations of VAR residuals

Autocorr.	$r_{M,t+1}$	$EVAR_{t+1}$	PE_{t+1}	$r_{Tbill,t+1}$	DEF_{t+1}	VS_{t+1}
unscaled						
$r_{M,t}$	-0.064	0.090	-0.058	-0.041	0.085	0.045
$EVAR_t$	0.073	-0.157	0.086	0.114	-0.188	-0.080
PE_t	-0.075	0.180	-0.141	-0.063	0.206	0.093
$r_{Tbill,t}$	0.002	0.016	-0.013	-0.139	-0.029	-0.057
DEF_t	0.132	-0.140	0.169	0.109	-0.289	-0.145
VS_t	0.021	-0.035	0.018	0.037	-0.085	-0.083
scaled						
$r_{M,t}$	0.005	0.042	0.001	-0.003	-0.001	-0.008
$EVAR_t$	0.061	-0.105	0.074	0.077	-0.124	-0.053
PE_t	-0.007	0.125	-0.072	-0.032	0.097	0.026
$r_{Tbill,t}$	-0.014	0.038	-0.028	-0.123	0.000	-0.036
DEF_t	0.080	-0.097	0.109	0.085	-0.202	-0.102
VS_t	0.020	-0.027	0.008	0.018	-0.076	-0.066

Table 3: Forecasting Long-Horizon Realized Variance

This Appendix Table studies the estimates of long-run variance implied by the VAR model of the paper. 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, the corresponding annualized long-horizon forecast implied from estimates of the VAR in the paper (VAR_{40}) as well as FI-GARCH (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. r_{Tbill} is the log three-month Treasury Bill yield. DEF is the default yield spread in percentage points, measured as the difference between the log yield on Moody's BAA bonds and the log yield on Moody's AAA bonds. VS is the small-stock value-spread, the difference in the log book-to-market ratios of small value and small growth stocks. 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 Appendix 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 Appendix 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	r_{Tbill}	DEF	VS	VAR_{40}	EG_{40}	FIG_{40}	PEO	$DEFO$	R^2%
1	-0.067	-0.007	0.099	0.023	0.067	0.012	0.003						56.66%
	[0.016]	[0.005]	[0.031]	[0.006]	[0.086]	[0.004]	[0.003]						
2	-0.009							1.017					49.83%
	[0.006]							[0.228]					
3	-0.067								1.458				40.01%
	[0.016]								[0.269]				
4	-0.006									0.987			37.31%
	[0.006]									[0.177]			
5	-0.105	-0.010	0.021	0.022	0.023	0.011	0.002		0.773				58.95%
	[0.022]	[0.006]	[0.023]	[0.005]	[0.081]	[0.003]	[0.004]		[0.277]				
6	-0.077	-0.008	-0.022	0.021	0.106	0.008	0.002			0.827			60.31%
	[0.015]	[0.006]	[0.021]	[0.006]	[0.080]	[0.003]	[0.003]			[0.225]			
7	-0.016							0.820		0.454			55.73%
	[0.006]							[0.211]		[0.221]			
8	-0.010					0.001		0.995					50.36%
	[0.006]					[0.003]		[0.232]					
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	0.046	0.046	0.045	0.045
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)	0.75	0.72	0.71	0.70

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
-0.024	1.552		-0.011	1.275		-0.006	1.176		-0.001	1.062	
[-1.22]	[3.97]		[-0.66]	[3.04]		[-0.35]	[2.73]		[-0.05]	[2.55]	
0.009		0.657	0.022		0.435	0.026		0.365	0.029		0.314
[0.62]		[3.30]	[2.34]		[2.20]	[2.47]		[2.27]	[2.26]		[1.96]
-0.022	1.255	0.194	-0.012	1.422	-0.099	-0.008	1.375	-0.127	-0.002	1.194	-0.083
[-1.08]	[2.11]	[0.66]	[-0.71]	[3.11]	[-0.69]	[-0.44]	[3.15]	[-0.93]	[-0.11]	[3.31]	[-0.65]

Table 4: Forecasting Long-Horizon Realized Variance: results across horizons

This Appendix Table reports the WLS parameter estimates of constrained regressions forecasting the annualized discounted sum of future $RVAR$ over the next h quarters $(4 * \sum_{k=1}^{h} \rho^{(k-1)} RVAR_{t+k} / \sum_{k=1}^{h} \rho^{(k-1)})$. The forecasting variables include the VAR state variables, the corresponding annualized long-horizon forecast (VAR_h) implied from estimates of the VAR in the paper as well as FIGARCH (FIG_h) and two-factor EGARCH (EG_h) 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. r^{Tbill} is the log three-month Treasury Bill yield. DEF is the default yield spread in percentage points, measured as the difference between the log yield on Moody's BAA bonds and the log yield on Moody's AAA bonds. VS is the small-stock value-spread, the difference in the log book-to-market ratios of small value and small growth stocks. Initial WLS weights are inversely proportional to the corresponding FIG_h long-horizon forecast except in those regressions involving VAR_h or EG_h forecasts, where the corresponding VAR_h or EG_h 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.

Varying the Horizon h in $(4 * \sum_{k=1}^{h} \rho^{(k-1)} RVAR_{t+k} / \sum_{k=1}^{h} \rho^{(k-1)})$

Constant	r_M	$RVAR$	PE	r_{Tbill}	DEF	VS	VAR_h	EG_h	FIG_h	$R^2\%$
				$h = 4$ (**1 year ahead**)						
-0.083	-0.025	0.198	0.027	-0.178	0.028	-0.001				47.20%
[0.024]	[0.023]	[0.101]	[0.009]	[0.195]	[0.010]	[0.010]				
-0.001							0.980			44.86%
[0.005]							[0.214]			
-0.007								1.054		47.11%
[0.004]								[0.172]		
-0.001									0.998	43.17%
[0.004]									[0.185]	
-0.064	-0.031	-0.141	0.020	-0.268	0.018	-0.003		0.897		54.22%
[0.019]	[0.019]	[0.092]	[0.007]	[0.165]	[0.007]	[0.009]		[0.211]		
-0.083	-0.029	-0.015	0.026	-0.150	0.022	-0.002			0.581	50.85%
[0.023]	[0.021]	[0.095]	[0.009]	[0.191]	[0.009]	[0.010]			[0.210]	
-0.007							0.708		0.492	50.04%
[0.004]							[0.206]		[0.186]	
-0.003					0.004		0.851			45.23%
[0.006]					[0.006]		[0.205]			
				$h = 8$ (**2 years ahead**)						
-0.101	-0.024	0.125	0.032	-0.137	0.027	0.003				44.21%
[0.028]	[0.017]	[0.082]	[0.011]	[0.206]	[0.011]	[0.010]				
-0.003							1.023			44.86%
[0.005]							[0.256]			
-0.013								1.024		37.58%
[0.007]								[0.217]		
0.001									0.936	32.24%
[0.006]									[0.234]	
-0.094	-0.025	-0.087	0.027	-0.186	0.019	0.000		0.717		48.07%
[0.027]	[0.017]	[0.082]	[0.010]	[0.189]	[0.009]	[0.010]		[0.223]		
-0.102	-0.027	0.019	0.032	-0.119	0.024	0.002			0.352	45.43%
[0.028]	[0.017]	[0.106]	[0.011]	[0.199]	[0.011]	[0.011]			[0.363]	
-0.008							0.866		0.326	44.48%
[0.006]							[0.262]		[0.271]	
-0.004					0.002		0.953			42.64%
[0.007]					[0.007]		[0.239]			
				$h = 20$ (**5 years ahead**)						
-0.078	-0.006	0.091	0.028	-0.120	0.020	-0.002				44.33%
[0.017]	[0.008]	[0.062]	[0.007]	[0.127]	[0.007]	[0.008]				
-0.004							0.932			39.58%
[0.005]							[0.243]			
-0.030								1.037		29.62%
[0.015]								[0.299]		
0.000									0.865	31.13%
[0.006]									[0.224]	
-0.087	-0.007	0.046	0.027	-0.130	0.018	-0.001		0.309		44.51%
[0.022]	[0.008]	[0.054]	[0.007]	[0.109]	[0.006]	[0.007]		[0.408]		
-0.080	-0.007	-0.011	0.027	-0.080	0.017	-0.002			0.471	45.63%
[0.007]	[0.008]	[0.043]	[0.007]	[0.129]	[0.007]	[0.008]			[0.363]	
-0.008							0.758		0.342	43.34%
[0.007]							0.178		0.283	
-0.005					0.002		0.895			39.88%
[0.006]					[0.004]		[0.188]			

Varying the Horizon h in $(4 * \sum_{k=1}^{h} \rho^{(k-1)} RVAR_{t+k} / \sum_{k=1}^{h} \rho^{(k-1)})$

Constant	r_M	$RVAR$	PE	r_{Tbill}	DEF	VS	VAR_h	EG_h	FIG_h	$R^2\%$
				$h = 60$ (**15 years ahead**)						
-0.060	-0.005	0.075	0.022	0.090	0.011	0.001				50.76%
[0.023]	[0.004]	[0.016]	[0.008]	[0.053]	[0.003]	[0.002]				
-0.012							1.056			41.54%
[0.006]							[0.224]			
-0.059								1.254		28.76%
[0.008]								[0.386]		
-0.003									0.812	30.86%
[0.041]									[0.210]	
-0.108	-0.007	0.023	0.024	0.061	0.010	0.000		0.765		52.63%
[0.041]	[0.004]	[0.022]	[0.008]	[0.056]	[0.002]	[0.002]		[0.442]		
-0.077	-0.007	-0.022	0.022	0.119	0.008	0.000			0.863	55.29%
[0.024]	[0.004]	[0.015]	[0.008]	[0.043]	[0.002]	[0.002]			[0.202]	
-0.016							0.857		0.343	45.71%
[0.005]							[0.252]		[0.264]	
-0.011					0.001		1.012			41.46%
[0.007]					[0.003]		[0.345]			

Table 5: Shocks to Short- and Long-run Expected Variance

This Appendix Table reports the correlation matrix for key shocks generated from our baseline VAR for the full sample (Panel A) as well as the early (Panel B) and modern (Panel C) subsamples. The shocks include N_V, the $EVAR$ shock generated from the second-stage of our VAR, the $RVAR$ generated from the first stage of our VAR, and the shock to r_M. The first two shocks represent innovations to short-run and long-run expected variance respectively.

Panel A: Full Sample

Correlations	N_V	$EVAR$ shock	$RVAR$ shock	r_m shock
N_V	1	0.66	0.45	-0.03
$EVAR$ shock	0.66	1	0.93	-0.51
$RVAR$ shock	0.45	0.93	1	-0.41
r_m shock	-0.03	-0.51	-0.41	1

Panel B: 1926-1963

Correlations	N_V	$EVAR$ shock	$RVAR$ shock	r_m shock
N_V	1	0.68	0.39	-0.31
$EVAR$ shock	0.68	1	0.88	-0.61
$RVAR$ shock	0.39	0.88	1	-0.36
r_m shock	-0.31	-0.61	-0.36	1

Panel C: 1963-2011

Correlations	N_V	$EVAR$ shock	$RVAR$ shock	r_m shock
N_V	1	0.65	0.50	0.27
$EVAR$ shock	0.65	1	0.96	-0.40
$RVAR$ shock	0.50	0.96	1	-0.47
r_m shock	0.27	-0.40	-0.47	1

Table 6: News Correlations and VAR specification

This Appendix Table reports correlations between news terms and returns innovations in different VAR specifications. Each specification 1-8 is obtained using different variables in the VAR. All VARs are estimated via OLS for the period 1930-2010 as in BKSY 2014. Each panel reports, for all 8 specifications of the state vector, a correlation using quarterly and yearly data. Two methods for computing RVAR are considered. ICSV uses squared daily returns, while BKSY uses squared monthly returns. The eight specifications vary the variables other than RVAR: 1) ICSV specification: same variables as our main VAR: Rm, PE, VS, DEF, TBILL. 2) BKSY specification: log dividend growth (Dd), log price-dividend ratio (PD), term spread (TS), long-term interest rate (LTR), default spread (DEF) as in BKSY 2014. 3) Use Rm (from ICSV) and PD, LTR, DEF, TS (from BKSY). 4) Use Rm, PE, DEF, TBILL (from ICSV) and TS (from BKSY). 5) Use Rm, PE, DEF, TBILL (from ICSV) and LTR (from BKSY). 6) Use Dd and PD (from BKSY) and TBILL, DEF, VS (from ICSV). 7) Use Dd, PD, LTR, DEF (from BKSY) and VS (from ICSV). 8) Use Dd, PD, TS, DEF (from BKSY) and VS (from ICSV). The correlation in bold in the first row corresponds to the baseline ICSV specification. The correlation in bold in the second row corresponds to the baseline BKSY specification.

	Full sample			Early period			Modern period		
	$\rho(N_{DR}, N_V)$			$\rho(N_{DR}, N_V)$			$\rho(N_{DR}, N_V)$		
	Quarterly	Yearly		Quarterly	Yearly		Quarterly	Yearly	
RVAR	ICSV	ICSV	BKSY	ICSV	ICSV	BKSY	ICSV	ICSV	BKSY
1	**-0.11**	-0.34	0.01	**0.26**	-0.25	0.14	**-0.41**	-0.43	-0.14
2	0.33	0.00	**0.47**	0.40	0.08	**0.54**	0.22	-0.08	**0.39**
3	0.51	0.08	0.58	0.58	0.15	0.69	0.40	0.00	0.45
4	0.16	-0.06	0.42	0.43	-0.02	0.44	-0.13	-0.10	0.38
5	0.16	-0.06	0.45	0.45	0.00	0.47	-0.14	-0.13	0.42
6	-0.03	-0.15	0.21	0.16	-0.09	0.31	-0.24	-0.20	0.11
7	-0.05	-0.08	0.32	0.15	-0.12	0.35	-0.29	-0.05	0.29
8	-0.03	-0.27	0.13	0.18	-0.15	0.29	-0.28	-0.37	0.00
	$\rho(N_{CF}, N_V)$			$\rho(N_{CF}, N_V)$			$\rho(N_{CF}, N_V)$		
	Quarterly	Yearly		Quarterly	Yearly		Quarterly	Yearly	
RVAR	ICSV	ICSV	BKSY	ICSV	ICSV	BKSY	ICSV	ICSV	BKSY
1	**-0.15**	-0.37	-0.58	**-0.14**	-0.36	-0.49	**-0.16**	-0.39	-0.68
2	-0.45	-0.55	**-0.63**	-0.60	-0.60	**-0.57**	-0.25	-0.51	**-0.71**
3	-0.36	-0.52	-0.75	-0.53	-0.60	-0.73	-0.17	-0.44	-0.78
4	0.11	-0.02	-0.25	0.02	-0.21	-0.32	0.21	0.19	-0.17
5	0.11	-0.08	-0.35	0.03	-0.23	-0.35	0.22	0.11	-0.36
6	-0.50	-0.60	-0.67	-0.66	-0.65	-0.66	-0.30	-0.55	-0.69
7	-0.53	-0.66	-0.69	-0.67	-0.72	-0.67	-0.36	-0.61	-0.72
8	-0.52	-0.62	-0.66	-0.67	-0.65	-0.62	-0.34	-0.59	-0.71
	$\rho(N_{CF}, N_{DR})$			$\rho(N_{CF}, N_{DR})$			$\rho(N_{CF}, N_{DR})$		
	Quarterly	Yearly		Quarterly	Yearly		Quarterly	Yearly	
RVAR	ICSV	ICSV	BKSY	ICSV	ICSV	BKSY	ICSV	ICSV	BKSY
1	**0.02**	-0.11	-0.10	**0.00**	-0.19	-0.18	**0.05**	-0.07	-0.05
2	-0.23	-0.36	**-0.35**	-0.15	-0.34	**-0.33**	-0.40	-0.38	**-0.37**
3	-0.22	-0.40	-0.37	-0.30	-0.41	-0.39	-0.11	-0.39	-0.35
4	-0.25	-0.59	-0.58	-0.14	-0.48	-0.49	-0.46	-0.72	-0.69
5	-0.26	-0.58	-0.57	-0.14	-0.44	-0.44	-0.48	-0.76	-0.73
6	-0.01	-0.20	-0.22	-0.01	-0.27	-0.27	-0.03	-0.14	-0.19
7	-0.07	-0.32	-0.30	-0.02	-0.21	-0.21	-0.15	-0.45	-0.42
8	-0.07	-0.15	-0.16	-0.02	-0.17	-0.16	-0.15	-0.13	-0.16
	$\rho(N_{CF} - N_{DR}, N_V)$			$\rho(N_{CF} - N_{DR}, N_V)$			$\rho(N_{CF} - N_{DR}, N_V)$		
	Quarterly	Yearly		Quarterly	Yearly		Quarterly	Yearly	
RVAR	ICSV	ICSV	BKSY	ICSV	ICSV	BKSY	ICSV	ICSV	BKSY
1	**0.03**	0.14	-0.24	**-0.29**	0.09	-0.30	**0.30**	0.18	-0.18
2	-0.50	-0.32	**-0.66**	-0.67	-0.40	**-0.68**	-0.28	-0.24	**-0.65**
3	-0.51	-0.36	-0.81	-0.67	-0.45	-0.85	-0.28	-0.27	-0.76
4	-0.09	0.04	-0.40	-0.35	-0.07	-0.46	0.18	0.14	-0.34
5	-0.09	0.01	-0.46	-0.37	-0.09	-0.50	0.19	0.13	-0.42
6	-0.37	-0.28	-0.56	-0.62	-0.35	-0.61	-0.08	-0.21	-0.51
7	-0.39	-0.39	-0.64	-0.63	-0.43	-0.67	-0.10	-0.34	-0.60
8	-0.37	-0.18	-0.50	-0.63	-0.32	-0.59	-0.07	-0.07	-0.41

Table 7: Average Excess Returns on Test Assets

This Appendix Table shows the average excess returns on 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), and the non-equity sample (Panel D) that includes the three equity factors of Fama and French (1993), the returns on high yield and investment grade bond portfolios, the five interest-rate-sorted currency portfolios from developed countries of Lustig, Roussanov, and Verdelhan (2011), the S&P 100 index straddle return from Coval and Shumway (2001), and three three-month variance forward positions with maturities of three, six, and nine months. The returns on the straddle and the VIX forwards are scaled down by a factor of 17 so that these assets have comparable average returns with the other test assets. "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 1 in ICSV, and on the market-return shock.

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

	1931:3-1963:2				
	Growth	2	3	4	Value
Small	3.45%	3.80%	6.13%	6.61%	7.35%
2	3.76%	5.36%	5.30%	5.62%	6.18%
3	4.42%	4.05%	4.79%	4.79%	5.73%
4	3.03%	3.87%	4.33%	4.51%	5.82%
Large	2.85%	2.53%	3.55%	3.95%	4.80%
	1963:3-2011:4				
	Growth	2	3	4	Value
Small	0.94%	2.52%	2.60%	3.12%	3.58%
2	1.49%	2.21%	2.88%	2.86%	3.12%
3	1.47%	2.31%	2.37%	2.65%	3.23%
4	1.76%	1.69%	2.08%	2.54%	2.53%
Large	1.28%	1.46%	1.32%	1.55%	1.63%

Panel B: 6 risk-sorted portfolios

	1931:3-1963:2		
	Lo $\widehat{b}_{r_M}$	2	Hi $\widehat{b}_{r_M}$
Lo $\widehat{b}_{VAR}$	2.74%	3.56%	4.48%
Hi $\widehat{b}_{VAR}$	2.74%	4.12%	4.67%
	1963:3-2011:4		
	Lo $\widehat{b}_{r_M}$	2	Hi $\widehat{b}_{r_M}$
Lo $\widehat{b}_{VAR}$	1.87%	2.19%	2.48%
Hi $\widehat{b}_{VAR}$	0.98%	1.29%	1.28%

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

1931:3-1963:2									
	Growth			2			Value		
	Low IVol	2	High IVol	Low IVol	2	High IVol	Low IVol	2	High IVol
Low $\widehat{\beta}_{\Delta VAR}$	2.40%	2.63%	3.88%	3.33%	4.45%	4.09%	4.96%	5.77%	6.41%
High $\widehat{\beta}_{\Delta VAR}$	2.72%	3.33%	3.14%	3.57%	4.35%	4.86%	5.02%	5.17%	6.03%
	$P1$: 2.54%			$P2$: Growth 0.95%, Value 1.22%			$P3$: 0.03%		
1963:3-2011:4									
	Growth			2			Value		
	Low IVol	2	High IVol	Low IVol	2	High IVol	Low IVol	2	High IVol
Low $\widehat{\beta}_{\Delta VAR}$	1.52%	1.39%	-0.22%	1.83%	2.18%	2.28%	2.02%	3.65%	3.51%
High $\widehat{\beta}_{\Delta VAR}$	1.06%	0.91%	-0.44%	1.47%	1.52%	1.18%	2.02%	3.03%	3.29%
	$P1$: 2.22%			$P2$: Growth -1.62%, Value 1.38%			$P3$: -0.46%		

Panel D: Non-equity sample

1998:1-2011:4				
RMRF	SMB	HML	IGRET	HYRET
1.50%	1.76%	1.37%	1.59%	1.75%
Low r*	2	3	4	High r*
0.97%	1.07%	1.31%	1.52%	2.22%
STRADDLE	VIXF0	VIXF1	VIXF2	
-0.48%	-0.37%	0.92%	1.20%	

Table 8: Cash-flow, Discount-rate, and Variance Betas

The table shows the estimated cash-flow ($\widehat{\beta}_{CF}$), discount-rate ($\widehat{\beta}_{DR}$), and variance betas ($\widehat{\beta}_{V}$) for the 25 ME- and BE/ME-sorted portfolios (Panels A and B) and six risk-sorted portfolios (Panels C and D) for the early (1931:3-1963:2) and modern (1963:3-2011:4) subsamples respectively as well as for the 18 BE/ME, IVol, and $\widehat{\beta}_{\Delta VAR}$-sorted portfolios in the modern period (Panel E) and the Fama-French factors $RMRF$, SMB, HML, high yield ($HYRET$) and investment grade ($IGRET$) bond portfolios, the five interest-rate-sorted portfolios of Lustig, Roussanov, and Verdelhan (2011) and the S&P 100 index straddle portfolio ($STRADDLE$) along with three VIX Forward positions (Panel F) over the common subperiod of 1998:1-2011:4. "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 2 of ICSV, and on the market-return shock. "Diff." is the difference between the extreme cells. Bootstrapped standard errors [in brackets] are conditional on the estimated news series. Estimates are based on quarterly data using weighted least squares where the weights are the same as those used to estimate the VAR.

25 ME- and BE/ME-sorted portfolios

Panel A: Early Period (1931:3-1963:2)

$\widehat{\beta}_{CF}$	Growth		2		3		4		Value		Diff	
Small	0.49	[0.13]	0.42	[0.11]	0.44	[0.11]	0.44	[0.10]	0.46	[0.10]	-0.04	[0.05]
2	0.30	[0.08]	0.36	[0.09]	0.37	[0.09]	0.39	[0.09]	0.42	[0.10]	0.12	[0.04]
3	0.32	[0.08]	0.29	[0.08]	0.34	[0.09]	0.33	[0.08]	0.47	[0.12]	0.15	[0.05]
4	0.26	[0.07]	0.28	[0.08]	0.31	[0.09]	0.35	[0.08]	0.44	[0.11]	0.18	[0.05]
Large	0.24	[0.07]	0.23	[0.07]	0.27	[0.09]	0.34	[0.10]	0.40	[0.29]	0.16	[0.04]
Diff	-0.26	[0.07]	-0.19	[0.05]	-0.17	[0.04]	-0.10	[0.03]	-0.06	[0.03]		

$\widehat{\beta}_{DR}$	Growth		2		3		4		Value		Diff	
Small	1.20	[0.15]	1.21	[0.16]	1.20	[0.17]	1.19	[0.17]	1.13	[0.17]	-0.07	[0.07]
2	0.87	[0.11]	1.03	[0.14]	1.01	[0.15]	0.99	[0.16]	1.14	[0.14]	0.27	[0.08]
3	0.95	[0.13]	0.81	[0.09]	0.97	[0.12]	0.93	[0.12]	1.22	[0.16]	0.27	[0.09]
4	0.67	[0.07]	0.81	[0.10]	0.85	[0.10]	0.93	[0.14]	1.24	[0.17]	0.58	[0.13]
Large	0.70	[0.08]	0.66	[0.08]	0.80	[0.12]	1.05	[0.16]	0.90	[0.12]	0.20	[0.13]
Diff	-0.50	[0.14]	-0.56	[0.11]	-0.40	[0.16]	-0.13	[0.13]	-0.23	[0.08]		

$\widehat{\beta}_{V}$	Growth		2		3		4		Value		Diff	
Small	-0.14	[0.05]	-0.14	[0.04]	-0.15	[0.05]	-0.14	[0.04]	-0.14	[0.04]	0.00	[0.02]
2	-0.08	[0.03]	-0.10	[0.03]	-0.10	[0.03]	-0.11	[0.03]	-0.14	[0.04]	-0.06	[0.02]
3	-0.09	[0.03]	-0.07	[0.02]	-0.09	[0.03]	-0.10	[0.03]	-0.14	[0.04]	-0.05	[0.02]
4	-0.04	[0.02]	-0.06	[0.02]	-0.08	[0.03]	-0.10	[0.04]	-0.15	[0.05]	-0.10	[0.03]
Large	-0.05	[0.02]	-0.05	[0.02]	-0.09	[0.04]	-0.12	[0.04]	-0.11	[0.03]	-0.07	[0.03]
Diff	0.09	[0.04]	0.09	[0.02]	0.06	[0.02]	0.02	[0.02]	0.03	[0.02]		

Panel B: Modern Period (1963:3-2011:4)

$\widehat{\beta}_{CF}$	Growth		2		3		4		Value		Diff	
Small	0.23	[0.06]	0.24	[0.05]	0.26	[0.05]	0.25	[0.04]	0.28	[0.05]	0.05	[0.04]
2	0.23	[0.06]	0.24	[0.05]	0.26	[0.05]	0.27	[0.05]	0.29	[0.05]	0.05	[0.04]
3	0.21	[0.05]	0.25	[0.05]	0.24	[0.05]	0.25	[0.05]	0.27	[0.05]	0.06	[0.03]
4	0.21	[0.05]	0.24	[0.04]	0.25	[0.04]	0.25	[0.04]	0.28	[0.05]	0.07	[0.03]
Large	0.15	[0.04]	0.20	[0.03]	0.18	[0.03]	0.20	[0.04]	0.20	[0.04]	0.05	[0.03]
Diff	-0.08	[0.04]	-0.04	[0.03]	-0.08	[0.03]	-0.05	[0.03]	-0.07	[0.03]		

$\widehat{\beta}_{DR}$	Growth		2		3		4		Value		Diff	
Small	1.30	[0.11]	1.05	[0.09]	0.87	[0.07]	0.81	[0.07]	0.86	[0.09]	-0.44	[0.08]
2	1.19	[0.09]	0.94	[0.08]	0.82	[0.07]	0.74	[0.07]	0.80	[0.08]	-0.39	[0.08]
3	1.11	[0.08]	0.87	[0.06]	0.73	[0.06]	0.70	[0.07]	0.69	[0.07]	-0.42	[0.08]
4	1.00	[0.07]	0.82	[0.06]	0.73	[0.07]	0.70	[0.07]	0.75	[0.07]	-0.26	[0.08]
Large	0.82	[0.05]	0.68	[0.04]	0.60	[0.05]	0.59	[0.07]	0.64	[0.06]	-0.18	[0.06]
Diff	-0.48	[0.10]	-0.37	[0.08]	-0.26	[0.06]	-0.22	[0.07]	-0.23	[0.08]		

$\widehat{\beta}_{V}$	Growth		2		3		4		Value		Diff	
Small	0.13	[0.07]	0.08	[0.06]	0.05	[0.05]	0.05	[0.05]	0.01	[0.07]	-0.13	[0.03]
2	0.14	[0.06]	0.08	[0.06]	0.05	[0.05]	0.04	[0.05]	0.03	[0.06]	-0.12	[0.02]
3	0.14	[0.06]	0.07	[0.05]	0.05	[0.05]	0.02	[0.05]	0.04	[0.04]	-0.10	[0.03]
4	0.13	[0.05]	0.07	[0.05]	0.03	[0.05]	0.02	[0.06]	0.01	[0.06]	-0.11	[0.02]
Large	0.09	[0.05]	0.07	[0.04]	0.03	[0.04]	0.02	[0.05]	0.02	[0.04]	-0.08	[0.02]
Diff	-0.04	[0.03]	-0.01	[0.03]	-0.02	[0.02]	-0.03	[0.02]	0.01	[0.03]		

Table 9: Cash-flow, Discount-rate, and Variance Betas: BE/ME, IVol, and Risk-sorted Portfolios

The table shows the estimated cash-flow ($\widehat{\beta}_{CF}$), discount-rate ($\widehat{\beta}_{DR}$), and variance betas ($\widehat{\beta}_{V}$) for the 18 BE/ME, IVol, and $\widehat{\beta}_{\Delta VAR}$-sorted portfolios (Panels A and B) for the early (1931:3-1963:2) and modern (1963:3-2011:4) subsamples respectively. $\widehat{\beta}_{\Delta VAR}$ is the 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 ICSV Table 2. $P1$ 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 using weighted least squares where the weights are the same as those used to estimate the VAR.

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

Panel A: Early Period (1931:3-1963:2)

$\widehat{\beta}_{CF}$	Growth						2						Value					
	Low IVol		2		High IVol		Low IVol		2		High IVol		Low IVol		2		High IVol	
Low $\widehat{\beta}_{\Delta VAR}$	0.24	[0.08]	0.31	[0.10]	0.38	[0.11]	0.25	[0.08]	0.39	[0.11]	0.43	[0.13]	0.37	[0.09]	0.47	[0.12]	0.47	[0.12]
High $\widehat{\beta}_{\Delta VAR}$	0.21	[0.06]	0.26	[0.07]	0.29	[0.08]	0.24	[0.07]	0.33	[0.10]	0.36	[0.09]	0.37	[0.10]	0.38	[0.10]	0.44	[0.10]
	$P1$: 0.31 [0.07]						$P2$: Growth 0.13 [0.03], Value 0.10 [0.03]						$P3$: -0.05 [0.02]					

$\widehat{\beta}_{DR}$	Growth						2						Value					
	Low IVol		2		High IVol		Low IVol		2		High IVol		Low IVol		2		High IVol	
Low $\widehat{\beta}_{\Delta VAR}$	0.69	[0.10]	0.89	[0.13]	1.07	[0.15]	0.86	[0.14]	1.03	[0.13]	1.21	[0.15]	0.96	[0.13]	1.18	[0.16]	1.16	[0.13]
High $\widehat{\beta}_{\Delta VAR}$	0.60	[0.07]	0.68	[0.08]	0.73	[0.08]	0.75	[0.11]	0.90	[0.12]	0.95	[0.13]	1.06	[0.15]	1.10	[0.15]	1.09	[0.16]
	$P1$: 1.05 [0.49]						$P2$: Growth 0.31 [0.10], Value 0.15 [0.05]						$P3$: -0.15 [0.04]					

$\widehat{\beta}_{V}$	Growth						2						Value					
	Low IVol		2		High IVol		Low IVol		2		High IVol		Low IVol		2		High IVol	
Low $\widehat{\beta}_{\Delta VAR}$	-0.04	[0.02]	-0.08	[0.03]	-0.11	[0.03]	-0.09	[0.04]	-0.10	[0.04]	-0.12	[0.03]	-0.10	[0.03]	-0.14	[0.05]	-0.13	[0.04]
High $\widehat{\beta}_{\Delta VAR}$	-0.04	[0.02]	-0.05	[0.02]	-0.06	[0.03]	-0.08	[0.04]	-0.09	[0.03]	-0.10	[0.04]	-0.12	[0.04]	-0.13	[0.04]	-0.14	[0.05]
	$P1$: -0.13 [0.06]						$P2$: Growth -0.05 [0.02], Value -0.02 [0.01]						$P3$: 0.02 [0.01]					

Panel B: Modern Period (1963:3-2011:4)

$\widehat{\beta}_{CF}$	Growth						2						Value					
	Low IVol		2		High IVol		Low IVol		2		High IVol		Low IVol		2		High IVol	
Low $\widehat{\beta}_{\Delta VAR}$	0.17	[0.03]	0.19	[0.05]	0.22	[0.06]	0.22	[0.04]	0.26	[0.05]	0.29	[0.06]	0.23	[0.04]	0.30	[0.06]	0.27	[0.06]
High $\widehat{\beta}_{\Delta VAR}$	0.17	[0.03]	0.16	[0.05]	0.19	[0.07]	0.18	[0.03]	0.22	[0.04]	0.25	[0.05]	0.20	[0.04]	0.25	[0.05]	0.32	[0.05]
	$P1$: 0.09 [0.02]						$P2$: Growth 0.03 [0.04], Value 0.08 [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 $\widehat{\beta}_{\Delta VAR}$	0.71	[0.05]	1.06	[0.10]	1.33	[0.12]	0.59	[0.05]	0.79	[0.07]	1.11	[0.09]	0.59	[0.06]	0.85	[0.09]	0.96	[0.12]
High $\widehat{\beta}_{\Delta VAR}$	0.74	[0.05]	1.07	[0.08]	1.37	[0.14]	0.57	[0.05]	0.86	[0.07]	1.08	[0.10]	0.56	[0.07]	0.88	[0.09]	1.05	[0.08]
	$P1$: -0.20 [0.05]						$P2$: Growth 0.58 [0.10], Value 0.44 [0.07]						$P3$: 0.00 [0.04]					

$\widehat{\beta}_{V}$	Growth						2						Value					
	Low IVol		2		High IVol		Low IVol		2		High IVol		Low IVol		2		High IVol	
Low $\widehat{\beta}_{\Delta VAR}$	0.07	[0.04]	0.08	[0.08]	0.14	[0.07]	0.03	[0.04]	0.02	[0.05]	0.07	[0.07]	0.01	[0.05]	0.01	[0.07]	-0.03	[0.09]
High $\widehat{\beta}_{\Delta VAR}$	0.10	[0.04]	0.12	[0.06]	0.13	[0.09]	0.04	[0.04]	0.06	[0.06]	0.09	[0.06]	0.05	[0.04]	0.02	[0.07]	0.07	[0.05]
	$P1$: -0.09 [0.02]						$P2$: Growth 0.05 [0.04], Value 0.00 [0.03]						$P3$: 0.03 [0.01]					

Table 10: Asset Pricing Tests: 25 Size and Book-to-Market Portfolios

The table reports GMM estimates of 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 for the early (Panel A: 1931:3-1963:2) and modern (Panel B: 1963:3-2011:4) subsamples. The test assets are 25 ME- and BE/ME-sorted portfolios and the T-bill with the market portfolio as the reference asset. The 5% critical value for the test of overidentifying restrictions is 36.5 in columns 1, 2, and 3; 35.2 in column 4; and 34.0 in column 5.

Parameter	CAPM	2-beta ICAPM	3-beta ICAPM	Constrained	Unrestricted
			Panel A: Early Period		
$\widehat{\beta}_{CF}$ premium (g_1)	0.040	0.102	0.082	0.040	0.074
Std. err.	(0.016)	(0.061)	(0.035)	(0.048)	(0.068)
$\widehat{\beta}_{DR}$ premium (g_2)	0.040	0.016	0.016	0.016	-0.003
Std. err.	(0.016)	0	0	(0.000)	(0.025)
$\widehat{\beta}_{VAR}$ premium (g_3)			-0.052	-0.157	-0.185
Std. err.			(0.066)	(0.162)	(0.177)
$\widehat{R^2}$	64%	66%	67%	68%	69%
J statistic	50.9	56.8	53.4	45.9	47.2
Implied γ	2.5	6.4	5.2	N/A	N/A
Implied ω	N/A	N/A	6.6	N/A	N/A
			Panel B: Modern Period		
$\widehat{\beta}_{CF}$ premium (g_1)	0.016	0.128	0.055	0.117	0.153
Std. err.	(0.010)	(0.047)	(0.000)	(0.050)	(0.045)
$\widehat{\beta}_{DR}$ premium (g_2)	0.016	0.008	0.008	0.008	-0.009
Std. err.	(0.010)	0	0	(0.000)	(0.015)
$\widehat{\beta}_{VAR}$ premium (g_3)			-0.096	-0.106	-0.033
Std. err.			(0.041)	(0.051)	(0.066)
$\widehat{R^2}$	-50%	45%	48%	76%	79%
J statistic	98.6	63.1	77.2	53.9	54.5
Implied γ	2.1	16.5	7.2	N/A	N/A
Implied ω	N/A	N/A	24.9	N/A	N/A

Table 11: Actual and Implied Consumption

This Appendix Table reports a comparison of aggregate consumption innovations, stockholders' and top stockholders' consumption innovations, and model-implied consumption innovations. We use yearly data between 1930 and 2011, except for stockholders' consumption, which is available only for the period 1982-2004. Real consumption growth is constructed from BEA data, stockholders' consumption is obtained from Malloy et al. (2009). Innovations in these variables are constructed by taking a residual of an AR(1) regression for each series. Panel A reports standard deviations of implied consumption innovations (for different values of ψ); aggregate consumption innovations (both in the full sample and in the subsample 1982-2004); and stockholders' and top stockholder's consumption innovations. Panel B reports the correlations between implied consumption and innovations in the three consumption series. Correlations with aggregate consumptions are reported both for the full sample (first column) and for the subsample for which stockholders' consumption data is available (1982-2004, second column).

Panel A: Standard Deviations

Δc (full sample)	0.019
Δc (82-04)	0.008
Δc (stockholders, 82-04)	0.041
Δc (top stockholders, 82-04)	0.174
Δc (implied, $\psi = 0.5$)	0.107
Δc (implied, $\psi = 1$)	0.178
Δc (implied, $\psi = 1.5$)	0.277

Panel B: Implied vs. actual consumption innovations

	Δc (full sample)	Δc (82-04)	Δc (stockh.,82-04)	Δc (top stockh.,82-04)
Δc (implied, $\psi = 0.5$)	0.32	-0.08	-0.02	0.21
Δc (implied, $\psi = 1$)	0.33	0.09	0.04	0.13
Δc (implied, $\psi = 1.5$)	0.30	0.13	0.05	0.08

Table 12: Correlation of Nrf with other News Terms

This Appendix Table reports the correlation matrix of the news terms (including implied risk-free rate news) with standard deviations on the diagonal.

	N_{cf}	N_{dr}	N_v	N_{rf}
N_{cf}	0.049	-0.041	-0.121	0.029
N_{dr}	-0.041	0.092	-0.034	0.851
N_v	-0.121	-0.034	0.025	-0.555
N_{rf}	0.029	0.851	-0.555	0.110

Table 13: Various Robustness Tests

This Appendix 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 ($D \equiv DEF$, $R \equiv r_{Tbill}$, $V \equiv VS$, $P \equiv PE$) 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_{Real}), 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 either out-of-sample versions of the model's news terms are used in the pricing tests or the VAR coefficients are allowed to vary across the early and modern subperiods. Panel F reports results using delevered market portfolios. Panel G 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 G also reports simple loadings of $RMRF$ and HML on $RVAR$ and the 15-year $FIGARCH$ variance forecast. Panel H reports time-series regressions explaining HML with the three news terms.

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

	None	D	D/R/V	**ALL**	P/D/V	P/D	P
$\widehat{\gamma}^{Max}$	4.5	3.2	3.1	**7.2**	6.7	8.9	14.8
				Early Period			
$RMRF\ \widehat{\beta}_V$	-0.03	-0.23	-0.20	**-0.03**	-0.03	-0.02	0.08
$SMB\ \widehat{\beta}_V$	-0.01	-0.08	-0.07	**-0.02**	-0.02	-0.02	0.03
$HML\ \widehat{\beta}_V$	0.00	-0.12	-0.13	**-0.06**	-0.06	-0.04	0.06
$\widehat{\gamma}$	2.4	2.4	2.5	**5.2**	5.0	5.4	8.1
$\widehat{\omega}$	3.1	3.1	3.6	**6.6**	5.7	5.0	12.4
$\widehat{R^2}$	66%	65%	66%	**67%**	68%	66%	54%
				Modern Period			
$RMRF\ \widehat{\beta}_V$	-0.11	-0.15	-0.07	**0.10**	0.11	0.07	0.00
$SMB\ \widehat{\beta}_V$	-0.03	-0.05	-0.02	**0.03**	0.03	0.01	0.00
$HML\ \widehat{\beta}_V$	0.00	-0.01	-0.11	**-0.11**	-0.11	-0.05	-0.02
$\widehat{\gamma}$	1.9	1.9	3.1	**7.2**	6.7	8.9	3.4
$\widehat{\omega}$	1.2	1.1	14.9	**24.9**	20.3	29.4	1.4
$\widehat{R^2}$	-55%	-55%	18%	**48%**	29%	-24%	-56%

Panel B: Results Using Different Estimation Methods

	All OLS	WLS 3	**WLS 5**	WLS 8	RVAR Weighted	Partial VAR
$\widehat{\gamma}^{Max}$	1.6	7.1	**7.2**	7.1	7.1	7.1
			Early Period			
$RMRF\ \widehat{\beta}_V$	-0.07	-0.05	**-0.03**	-0.01	-0.01	0.01
$SMB\ \widehat{\beta}_V$	-0.03	-0.02	**-0.02**	-0.02	-0.02	-0.01
$HML\ \widehat{\beta}_V$	-0.08	-0.07	**-0.06**	-0.05	-0.05	-0.03
$\widehat{\gamma}$	1.6	5.2	**5.2**	5.3	4.8	3.9
$\widehat{\omega}$	2.3	6.5	**6.6**	6.9	5.9	15.9
$\widehat{R^2}$	28%	67%	**67%**	67%	70%	62%
			Modern Period			
$RMRF\ \widehat{\beta}_V$	0.08	0.09	**0.10**	0.11	0.10	0.11
$SMB\ \widehat{\beta}_V$	0.02	0.02	**0.03**	0.03	0.03	0.03
$HML\ \widehat{\beta}_V$	-0.10	-0.10	**-0.11**	-0.11	-0.13	-0.12
$\widehat{\gamma}$	1.6	7.1	**7.2**	7.1	5.8	2.8
$\widehat{\omega}$	2.3	24.4	**24.9**	25.3	12.8	2.8
$\widehat{R^2}$	-44%	46%	**48%**	49%	-9%	-207%

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
$\widehat{\gamma}^{Max}$	7.2	6.8	4.9	**5.4**	5.8	5.9
			Early Period			
$RMRF\ \widehat{\beta}_V$	**-0.03**	0.16	-0.19	-0.20	0.11	0.14
$SMB\ \widehat{\beta}_V$	**-0.02**	0.07	-0.08	-0.06	0.04	0.04
$HML\ \widehat{\beta}_V$	**-0.06**	0.06	-0.15	-0.13	0.01	0.09
$\widehat{\gamma}$	**5.2**	7.2	4.0	4.5	5.4	5.8
$\widehat{\omega}$	**6.6**	24.9	4.0	4.1	11.3	15.4
$\widehat{R^2}$	**67%**	50%	67%	63%	31%	60%
			Modern Period			
$RMRF\ \widehat{\beta}_V$	**0.10**	0.08	0.03	0.03	0.07	0.04
$SMB\ \widehat{\beta}_V$	**0.03**	-0.01	0.00	0.00	-0.02	-0.03
$HML\ \widehat{\beta}_V$	**-0.11**	-0.06	-0.13	-0.15	-0.06	-0.08
$\widehat{\gamma}$	**7.2**	0.0	4.9	5.4	5.4	5.8
$\widehat{\omega}$	**24.9**	0.3	12.7	11.3	11.3	15.4
$\widehat{R^2}$	**48%**	-48%	43%	43%	-24%	-41%

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

	PE	$PE_{\mathrm{Re}al}$	PD	$INFL$	CAY	FIG	$TYVOL$
$\widehat{\gamma}^{Max}$	**7.2**	9.6	4.6	9.3	9.3	9.3	7.2
			Early Period				
$RMRF\ \widehat{\beta}_V$	**-0.03**	0.03	-0.13	0.02	0.07	-0.02	-0.02
$SMB\ \widehat{\beta}_V$	**-0.02**	-0.01	-0.05	-0.01	0.01	-0.02	-0.02
$HML\ \widehat{\beta}_V$	**-0.06**	-0.02	-0.11	-0.03	0.00	-0.02	-0.05
$\widehat{\gamma}$	**5.2**	6.0	3.2	6.1	11.4	4.6	5.1
$\widehat{\omega}$	**6.6**	6.5	3.2	6.8	23.0	5.6	6.4
$\widehat{R^2}$	**67%**	67%	69%	66%	-689%	70%	67%
			Modern Period				
$RMRF\ \widehat{\beta}_V$	**0.10**	0.13	-0.01	0.14	0.06	0.11	0.09
$SMB\ \widehat{\beta}_V$	**0.03**	0.03	-0.01	0.03	0.01	0.03	0.02
$HML\ \widehat{\beta}_V$	**-0.11**	-0.09	-0.09	-0.09	-0.05	-0.08	-0.11
$\widehat{\gamma}$	**7.2**	9.6	4.6	9.3	15.9	5.3	7.2
$\widehat{\omega}$	**24.9**	31.4	18.3	29.4	58.2	9.7	24.9
$\widehat{R^2}$	**48%**	38%	14%	38%	26%	-58%	51%

Panel E: Modern Period Pricing

	Full Sample	Out of Sample	Structural Break
$RMRF\ \widehat{\beta}_V$	**0.10**	-0.10	0.08
$SMB\ \widehat{\beta}_V$	**0.03**	-0.04	0.02
$HML\ \widehat{\beta}_V$	**-0.11**	-0.10	-0.12
$\widehat{\gamma}$	**7.2**	6.3	4.6
$\widehat{\omega}$	**24.9**	11.9	22.6
$\widehat{R^2}$	**48%**	77%	48%

Panel F: Delevered Market

Equity %	**100%**	80%	60%
$\widehat{\gamma}^{Max}_{Exact}$	**7.2**	10.2	14.5
Early Period			
$RMRF\ \widehat{\beta}_V$	**-0.03**	-0.03	-0.03
$SMB\ \widehat{\beta}_V$	**-0.02**	-0.02	-0.02
$HML\ \widehat{\beta}_V$	**-0.06**	-0.06	-0.05
$\widehat{\gamma}$	**5.2**	7.6	11.3
$\widehat{\omega}$	**6.6**	15.7	49.1
$\widehat{R^2}$	**67%**	68%	69%
Modern Period			
$RMRF\ \widehat{\beta}_V$	**0.10**	0.10	0.10
$SMB\ \widehat{\beta}_V$	**0.03**	0.03	0.02
$HML\ \widehat{\beta}_V$	**-0.11**	-0.11	-0.11
$\widehat{\gamma}$	**7.2**	10.2	14.5
$\widehat{\omega}$	**24.9**	50.1	128.9
$\widehat{R^2}$	**48%**	45%	35%

Panel G: 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
$\widehat{\beta}_{\lambda_V^1 r_M\ Shock}$	0.01	0.01	0.00	0.00
$\widehat{\beta}_{\lambda_V^2 EVAR\ Shock}$	-0.06	-0.06	-0.03	-0.03
$\widehat{\beta}_{\lambda_V^3 PE\ Shock}$	0.14	0.14	0.06	0.07
$\widehat{\beta}_{\lambda_V^4 r^{Tbill}\ Shock}$	0.00	0.00	0.00	0.00
$\widehat{\beta}_{\lambda_V^5 DEF\ Shock}$	-0.13	-0.12	-0.09	-0.08
$\widehat{\beta}_{\lambda_V^6 VS\ Shock}$	-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
$\widehat{\beta}_{\lambda_V^1 r_M\ Shock}$	0.01	0.01	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	0.13	-0.02	-0.02
$\widehat{\beta}_{\lambda_V^4 r^{Tbill}\ Shock}$	0.00	0.00	0.00	0.00
$\widehat{\beta}_{\lambda_V^5 DEF\ Shock}$	-0.01	-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 H: Time-series Regressions explaining HML

Early Period				
	(1)		(2)	
Intercept	0.01	1.39	0.01	2.30
N_{CF}			0.39	3.68
$-N_{DR}$			0.40	6.67
N_V	-1.96	-6.46	-1.30	-4.96
$\widehat{R^2}$	24%		49%	

Modern Period				
	(1)		(2)	
Intercept	0.01	2.28	0.01	2.67
N_{CF}			0.24	2.70
$-N_{DR}$			-0.23	-4.92
N_V	-1.05	-7.34	-0.82	-5.84
$\widehat{R^2}$	22%		32%	

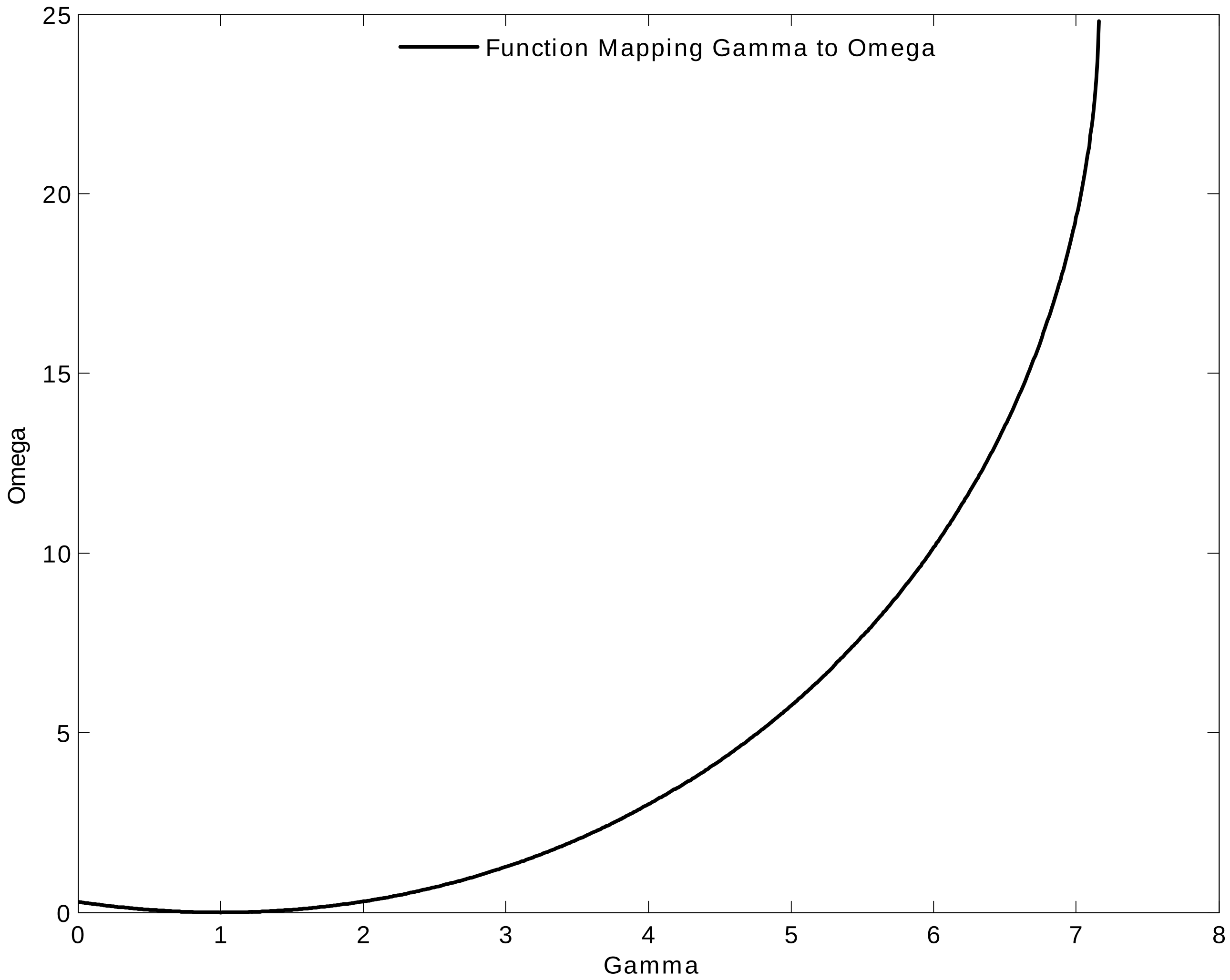

Figure 1: This figure graphs the relation between the parameter γ and the parameter ω described by equation (22). These functions depend on the loglinearization parameter ρ, set to 0.95 per year and the empirically estimated VAR parameters of ICSV Table 1. γ is the investor's risk aversion while ω is the sensitivity of news about risk, N_{RISK}, to news about market variance, N_V.

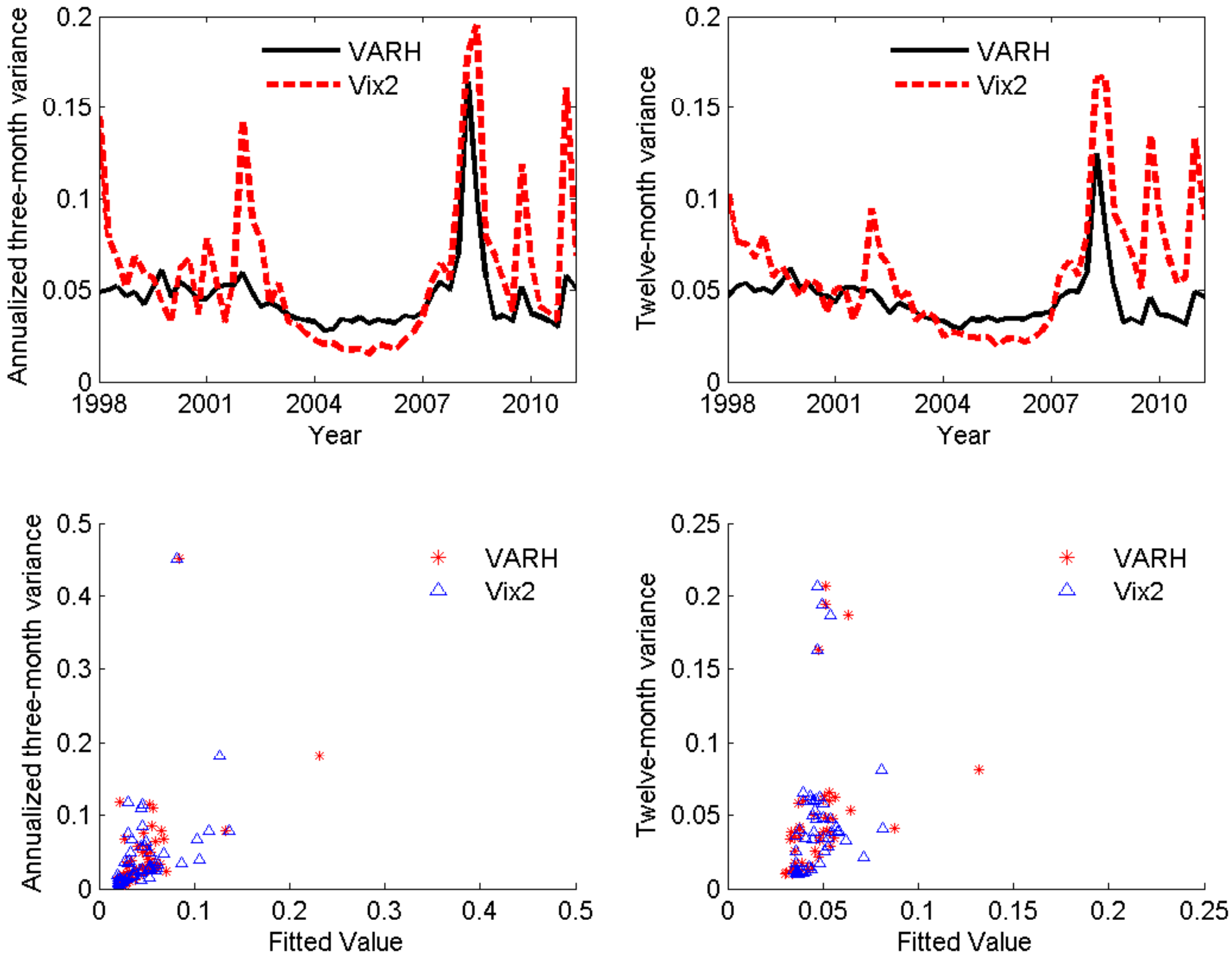

Figure 2: The top two diagrams correspond to forecasts of three-month (top left panel) and twelve-month (top right panel) variance from the VAR (VAR_h, solid black line) and from the option market (VIX_h^2, dashed red line). The bottom two diagrams correspond to scatter plots of three-month (bottom left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

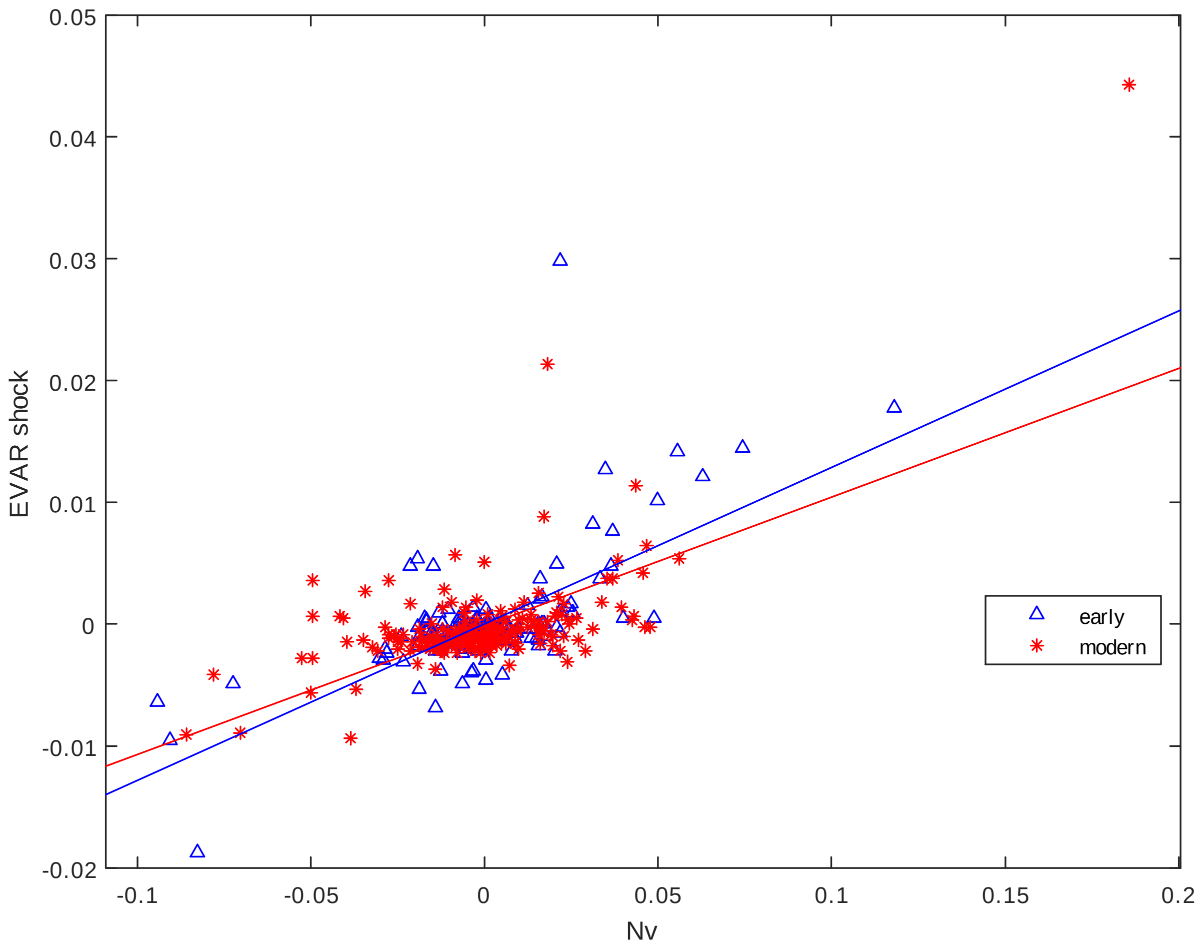

Figure 3: This figure plots shocks to $EVAR$ against volatility news, N_V. Observations from the early period are denoted with blue triangles while observations from the modern period are denoted with red asterisks.

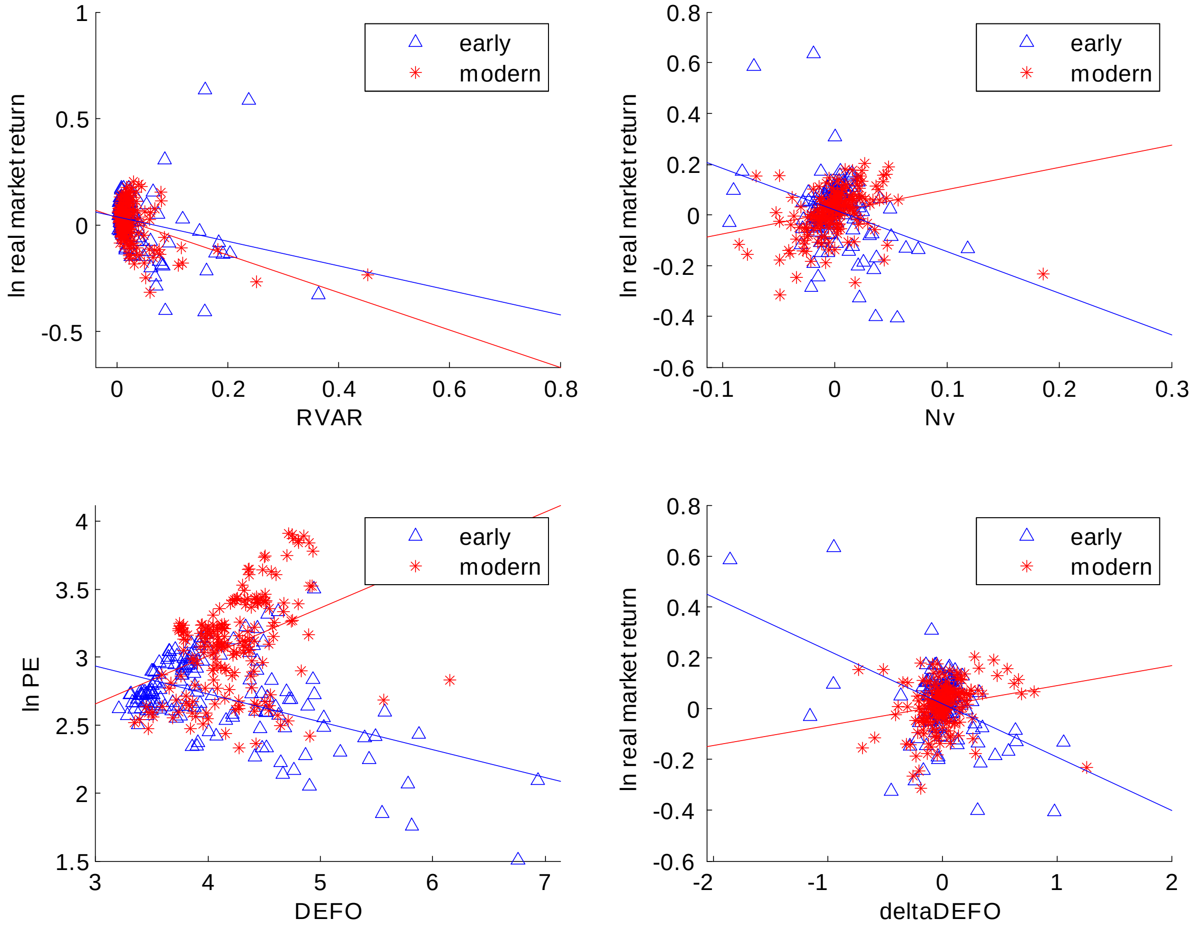

Figure 4: The top left portion of the figure plots the market return against $RVAR$. The top right portion of the figure plots the market return against volatility news, N_V. The bottom left of the figure plots PE against $DEFO$ (DEF orthogonalized to PE). The bottom right of the figure plots market 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.

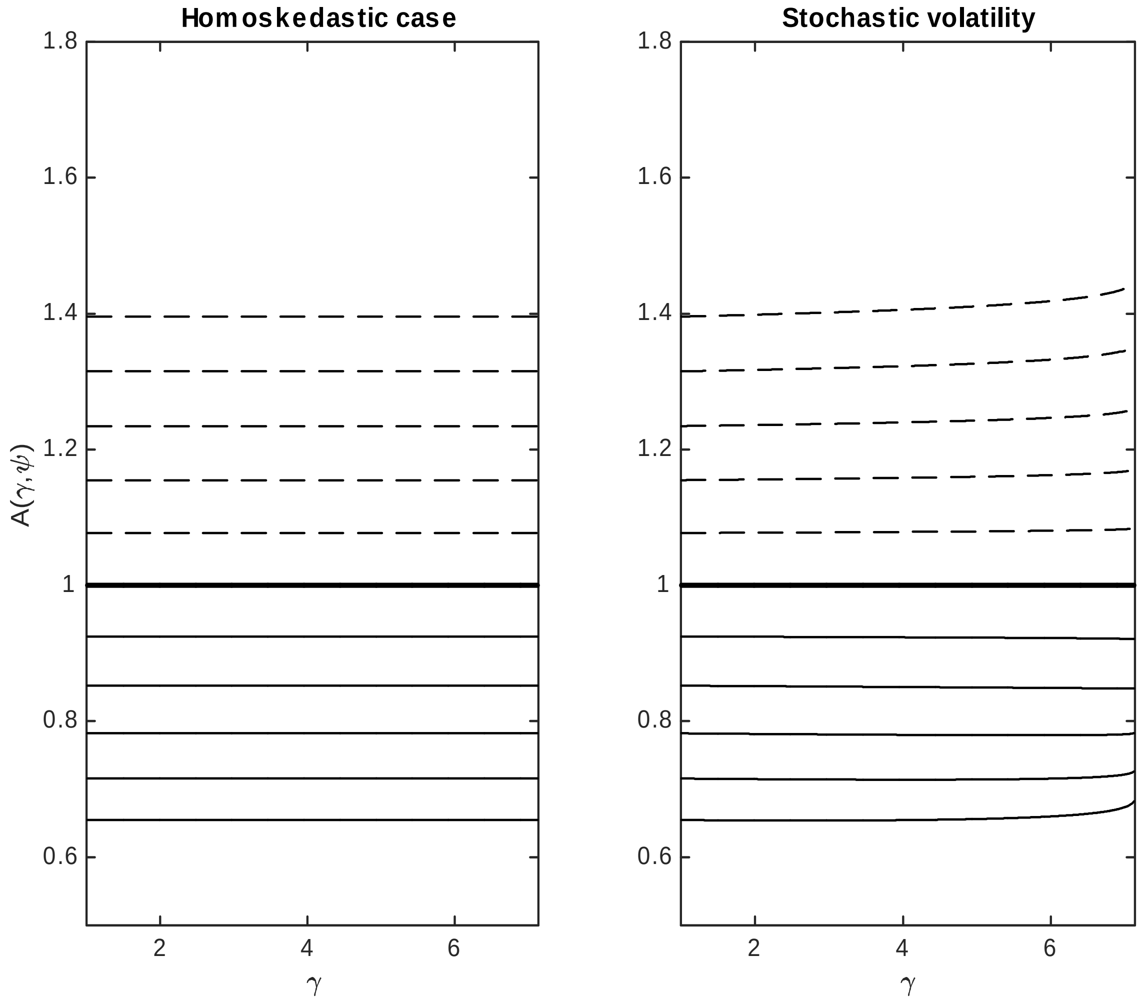

Figure 5: This figure plots plots the coefficient $A(\gamma, \psi)$ relating the conditional volatility of consumption growth to the volatility of returns for different values of γ and ψ for the homoskedastic case (left panel) and for the heteroskedastic case (right panel), where $A(\gamma, \psi)$ is a function of the variances and covariances of the *scaled* residuals u_{t+1}. In each panel, we plot $A(\gamma, \psi)$ as γ varies between 1 and the maximum possible value, for different values of ψ. Each line corresponds to a different ψ, beginning with the bottommost solid line line (ψ=0.5), incrementing ψ by 0.1 until ending with the topmost dashed line (ψ=1.5).

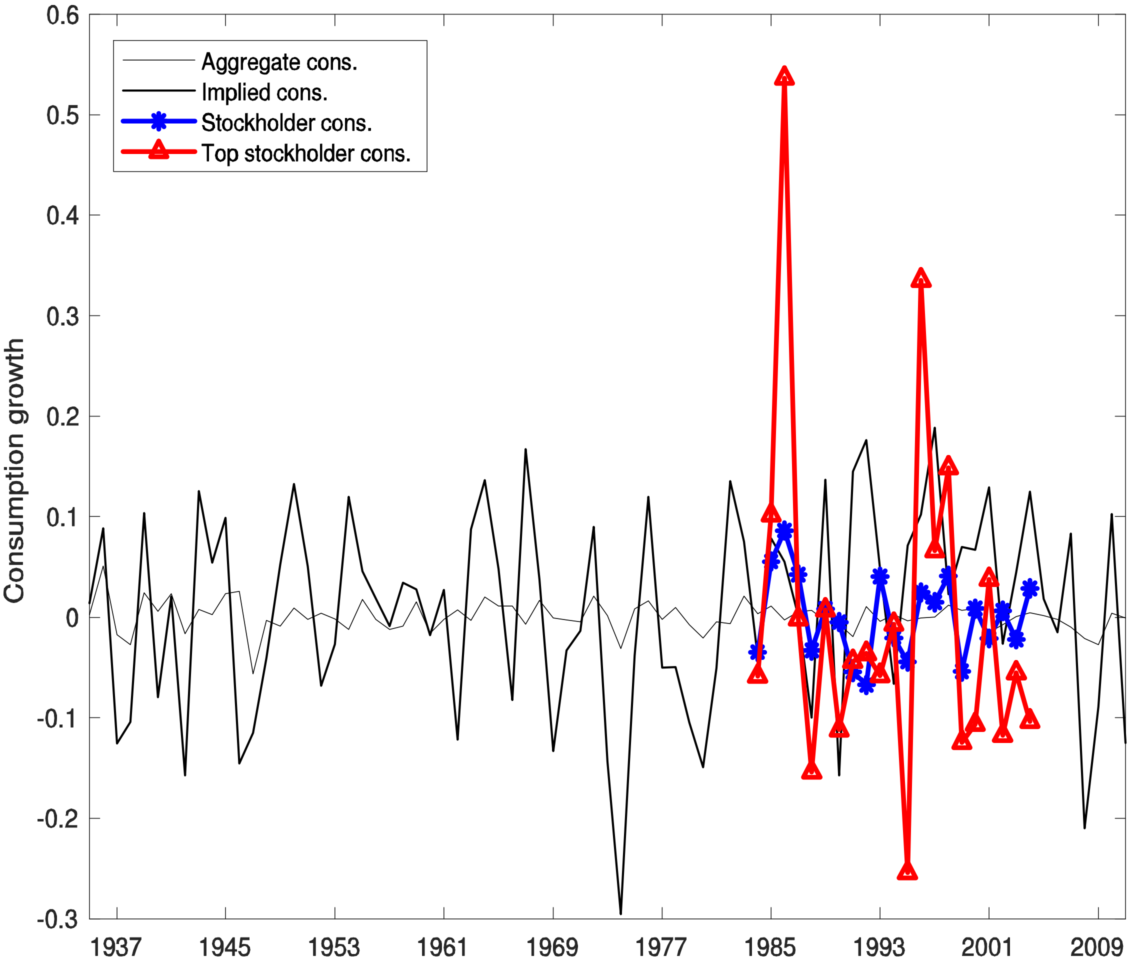

Figure 6: We plot the time series of aggregate consumption growth, stockholders' and top stockholders' consumption growth from Malloy et al. (2009), and implied consumption growth, obtained setting $\psi = 0.5$.

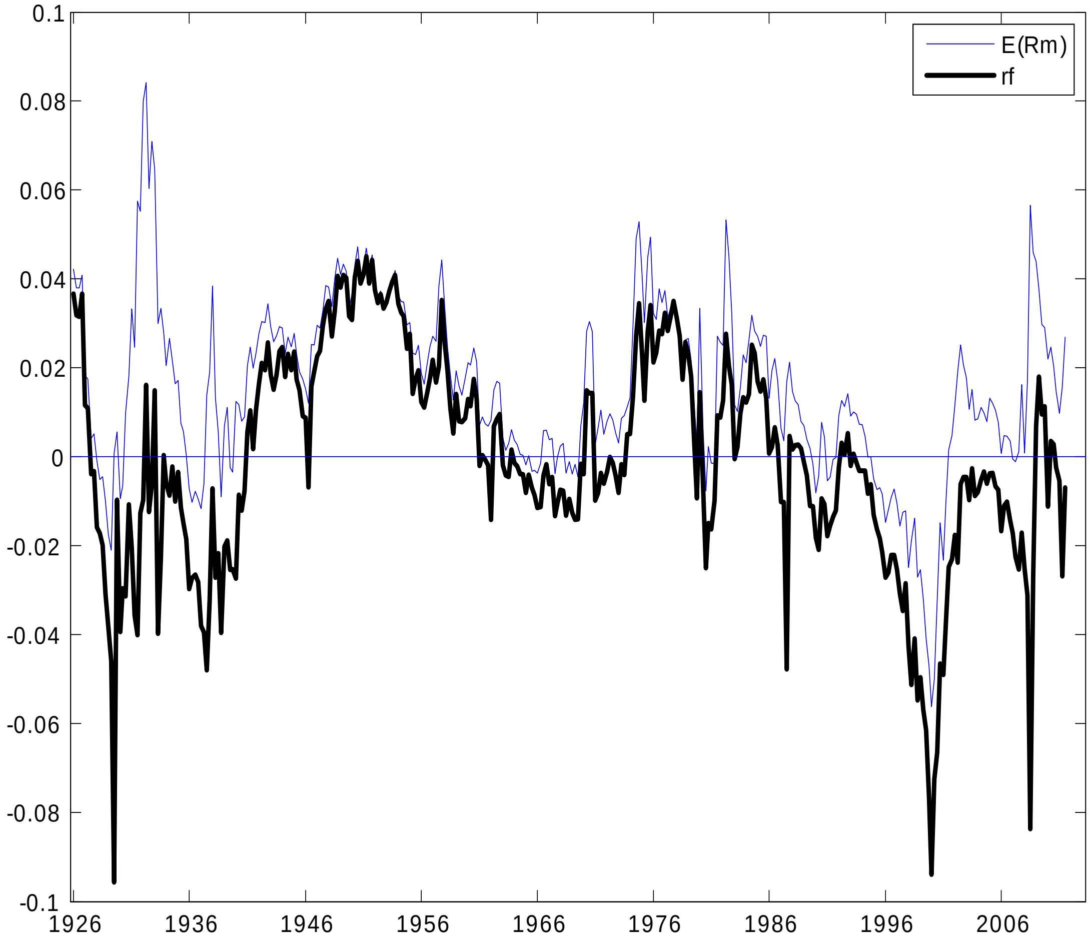

Figure 7: The figure reports the time series of $E_t r^M_{t+1}$ in the estimated model as well as the implied risk-free rate, r^f_t.

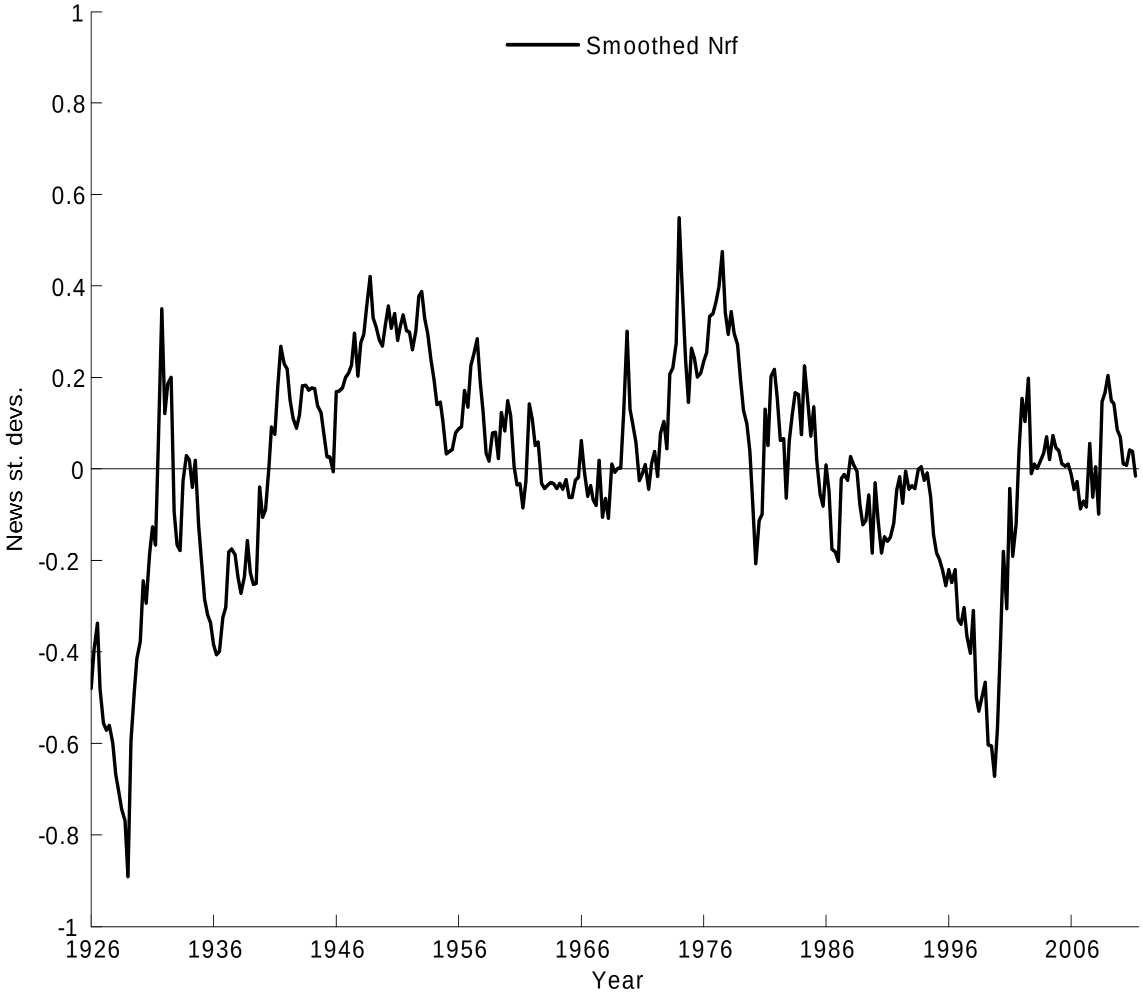

Figure 8: This figure plots implied risk-free rate news. The series is first normalized by dividing by its standard deviation, and then smoothed with a trailing exponentially-weighted moving average where the decay parameter is set to 0.08 per quarter. The smoothed normalized news series is generated as $MA_t(N) = 0.08N_t + (1-0.08)MA_{t-1}(N)$. The parameter implies a half-life of two years. The sample period is 1926:2-2011:4.