Skewness in Expected Macro Fundamentals and the Predictability of Equity Returns: Evidence and Theory

Riccardo Colacito Eric Ghysels Jinghan Meng*

Abstract

We show that introducing time-varying skewness in the distribution of expected growth prospects in an otherwise standard endowment economy can up to double the model implied equity Sharpe ratios, and produce a substantial amount of fluctuation in equity risk premia. Looking at the Livingston Survey, we document that the first and third cross-sectional moments of the distribution of GDP growth rates made by professional forecasters can predict equity excess returns, a finding which is consistent with our consumption based asset pricing model.

JEL classification: C22; C58; G12.
First Draft: August 10, 2012. *This draft*: October 23, 2012.

Most recent draft available at: http://www.unc.edu/~colacitr/Research/CGM.pdf

*All authors are affiliated with the University of North Carolina at Chapel Hill. We thank seminars participants at the University of North Carolina at Chapel Hill.

1 Introduction

Each month a large number of forecasts about expected growth prospects of the economy is made available to the public. A lot of attention is typically devoted to the average of all these forecasts, sometimes called the consensus forecast. Indeed ex post assessments of the stock market reactions to the release of official economic data is made relative to the discrepancy between actual data and ex-ante consensus forecast. The consensus forecast is inherently an *average of the average* forecasts, as the entire distribution of the forecasts is generally not available for each analyst.

In this paper, we put ourselves in the position of an investor that looks at the entire distribution of analysts' average forecasts. This is a particularly relevant exercise because, if on the one hand the consensus forecast may provide a reliable prediction for the near term, on the other hand the entire distribution of forecasts may contain useful information to assess more precisely the medium and longer-term growth prospects of the economy.

Using the Livingston Survey dataset, we find that the degree of asymmetry of the distribution of professional forecasters helps in predicting future expected growth rates. This is true even without including the period of the financial crisis of 2008. We also find that the degree of asymmetry is moderately persistent. We explore the asset pricing implication of a model in which expected growth features a time-varying degree of skewness.

Motivated by the empirical evidence concerning the time-varying shape of expected growth prospects, we investigate the importance of modeling time-varying skewness in the context of a consumption based asset pricing model. We follow Bansal and Yaron (2004) by assuming that investors order consumption profiles using Epstein and Zin (1989) preferences. This means that agents care about the temporal distribution of risk. In particular, we show that this type of investors not only likes high expected utility levels, but also dislikes uncertainty and negative asymmetry about her future utility. We explicitly model the expected growth rate of consumption as following a skew-normal distribution with time-varying parameters. First introduced by Azzalini (1985), the skew-normal provides a convenient way of modeling asymmetric distributions, as the first three conditional moments are available in closed form. Furthermore, we can easily incorporate the empirical finding that the cross-sectional mean, variance, and skewness of the distribution of forecasts follow autoregressive processes and that the cross-sectional skewness appears to have predictive power for the conditional mean. We show that the introduction of skewness can: i) up to double equity risk premia, and ii) produce a substantial amount of time variation in conditional risk premia.

Given the mean forecast for each analyst, investors can postulate a transition model for the distribution of the conditional mean of GDP growth. The fact that we focus on the cross-section of analyst's mean forecasts highlights a subtle but crucial difference between the contributions of our paper and the rare events literature. The case of the financial crisis of 2008 provides a useful example. During 2008-2009 there was a lot of discussion regarding a rise in equity premia as a compensation for the increased probability of a catastrophic event in the economy. Looking at the distribution of analysts' forecasts, this was reflected in a large drop in the consensus forecast, which ranged between -0.2 and 0.8 during the recession. The skewness of the distribution of average forecasts, however, was mostly positive during this period (on average it was 0.47). Quite clearly, the extra equity premium that we are capturing in this paper is not coming from an increase in negative skewness during a recession. Instead, it is coming from the fact that negative skewness today predicts that the future revisions of the average growth rate will be more pessimistic. Indeed, the skewness of the forecasts' distribution was very negative heading into the recession (-2.10 in the first semester of 2007) at a time when the average forecast was a solid 2%. The important message for macro-finance models is that skewness matters above and beyond its role as a signal for an increased probability of a catastrophic event. Skewness matters because it is the indication of possibly long-lasting movements in the expected growth rate of the economy. This uncertainty matters in general in the long-run risks framework and it is reinforced here by the presence of time-varying skewness.

Using the Livingston dataset, we confront our model with the data. We document that the cross-sectional moments of the distribution of professional forecasters' expected GDP growth help predict future equity excess returns. In particular, we show that the first and third cross-sectional moments have economically and statistically significant predictive power, as larger mean and more positive skewness predict lower equity returns going forward. This empirical result remains even after one controls for standard predictors such as cay, dividend yields, and default premia (see Goyal and Welsch (2008) for a comprehensive study). These findings are consistent with our consumption based asset pricing model.

This paper is related to several strands of the literature. An extensive literature has documented the predictability of equity excess returns at various horizons (see again inter alia Goyal and Welsch (2008)). Campbell and Diebold (2009) have provided evidence in support of the predictive power of the consensus forecast for subsequent stock market returns. We extend their findings and show that the degree of asymmetry can also help explain equity returns going forward. This paper builds on the recent literature on long-run risks by showing that the introduction of skewness in the dynamics of a small, but highly persistent predictive component of consumption growth can further amplify the ability of equilibrium models of consumption to account for asset pricing phenomena. Furthermore, there is a considerable literature on asset pricing models with investors who take into account higher moments (beyond variances) in *asset returns*. Arditti (1967), Rubinstein (1973), Kraus and Litzenberger (1976), Harvey and Siddique (2000) developed some of the early models of expected returns which incorporate the higher moments of individual securities that co-move with the aggregate market portfolio. Subsequently, empirical work provided supporting evidence that higher moments of the return distribution are important in pricing securities (see e.g. Harvey and Siddique (2000), and most recently Chang, Christoffersen and Jacobs (2012), Conrad, Dittmar and Ghysels (2012), among others).

There have also been several other attempts to find structural asset pricing interpretations of skewness. For example Damodaran (1985) suggests skewed distributions of asset returns are caused by investors reacting asymmetrically to good and bad company news. Chen, Hong and Stein (2001) argue that differences of opinion among investors combined with short-sale constraints generate skewed returns. Chabi-Yo, Ghysels, and Renault (2010) show that allowing for heterogeneity in investors' preferences and beliefs can give rise to additional factors related to skewness and kurtosis in the pricing of nonlinear risks, whereas Mitton and Vorkink (2007) show that allowing for heterogeneity in investors' preferences for skewness can also lead to right-skewed securities having higher prices.

The paper is organized as follows. Section 2 documents the time series properties of the cross-sectional moments of the distribution of expected real GDP growth rates. Section 3 describes the types of preferences that are employed throughout our theoretical analysis, as well as the postulated dynamics of consumption growth. Section 4 reports the results from a calibrated version of the proposed economy, and section 5 details its asset pricing implications. Section 6 confronts the empirical predictions of the model with the data, and section 7 concludes the paper.

2 Time series properties of the cross-section of expected GDP growth

Multiple forecasts are commonly available for key economic variables, as different professional forecasters may disagree about the outlook of the economy, or simply because different forecasting models are employed in this task. In this section, we put

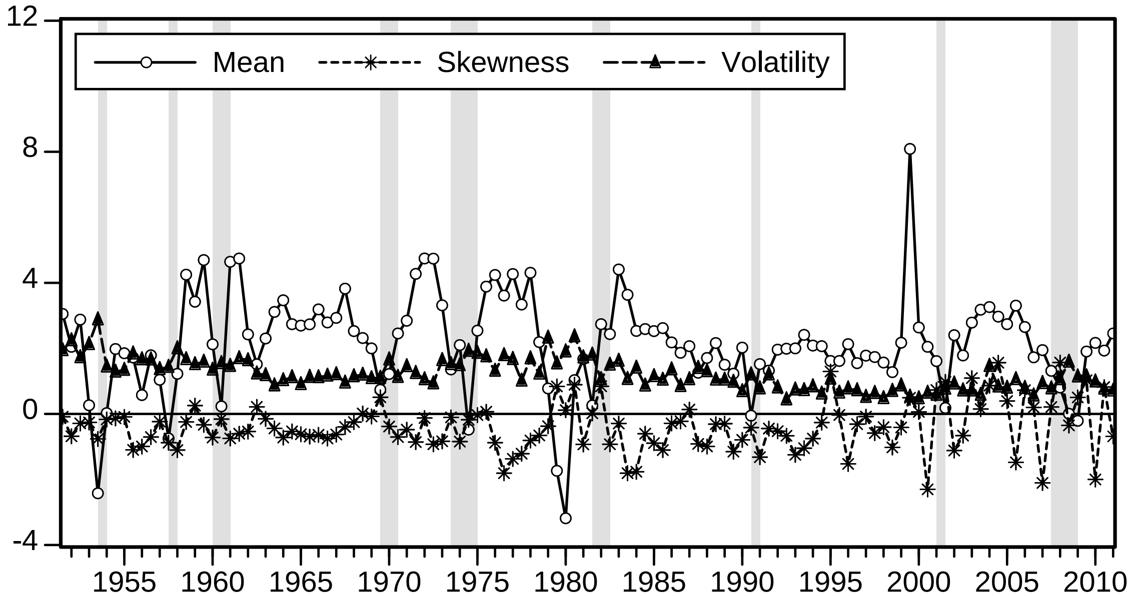

FIG. 1 - Time series for the first three cross-sectional moments of the distribution of expected real GDP growth. The series are constructed using one semester ahead real GDP growth forecasts from the Livingston dataset from 1951:1 to 2011:2. The vertical grey bars represent recessions according to the National Bureau of Economic Research.

ourselves in the position of an investor who looks at the entire cross-sectional distribution of these forecasts at each point in time. We document that this distribution features time-varying mean, volatility, and skewness.

Dataset. We construct the time-series of cross-sectional measures of mean, dispersion, and asymmetry of GDP growth expectations using the Livingston Survey. This survey was started in 1946 and it is the oldest continuous survey of economists' expectations. It summarizes the forecasts of economists from industry, government, banking, and academia. The Federal Reserve Bank of Philadelphia took responsibility for the survey in 1990. Every June and December, the Livingston Survey asks participants to forecast a set of key macroeconomic variables, including real and nominal GDP. Survey participants are asked to provide forecasts for these variables for the end of the current month, six months ahead, and 12 months ahead. For each date we have a cross-section of up to 50 forecasts. Our interest in this specific survey is motivated by the fact that it spans the longest time period, an appealing feature since we are trying to capture the properties of a slowly moving component of GDP growth.[1]

Time varying moments. Figure 1 reports the time-series for the first three cross-sectional moments of the distribution of expected GDP growth.[2] The figure shows that these moments are varying over time. While average expected GDP growth is on average positive, the skewness is negative in the most of the occasions. The dispersion of the forecasts appears to be very persistent. Quite interestingly, the three moments seem to be almost uncorrelated with one another. This suggests that the asymmetry of the distribution of forecasts may contain additional information about the risk factors in the economy. A qualitative finding that can be appreciated from looking at Figure 1 is that skewness tends to turn more negative right before the beginning of recessions, even at times when the mean forecast would otherwise suggest normal growth rates. This effect is particularly apparent for the last two recessions. We investigate this empirical regularity in greater detail in the last section of the paper.

Time series regressions. Table 1 reports some additional information about the time series properties of the cross-sectional moments of the distribution of average forecasts. In Panel A, we estimate three separate AR(1) processes for the mean, the volatility, and the third centered moment to the power of $1/3$. We choose to focus on this specific power of the third moment, because the model that we propose in the later sections directly imposes restrictions on its dynamics. Our time series estimates suggest that all three moments feature statistically significant first order autocorrelations. The persistence appears to be more pronounced for the first two moments.

[1] Since 1968 a richer data set that includes individual density forecasts is available. To motivate the theory and build our empirical models we opted for the longer time series with only the individual point forecasts.

[2] Volatility and in particular skewness estimates may be sensitive to outliers. We therfore also considered the quantile-based measures of volatility and skewness (see e.g. Conrad, Dittmar and Ghysels (2012) for details about such measures) to control for the effect of outliers. The results are similar as those depicted in Figure 1 and available on request.

TABLE 1
TIME SERIES PROPERTIES OF CROSS-SECTIONAL MOMENTS

Panel A	Mean	Volatility	Third Moment$^{1/3}$
Lagged Mean	0.556	–	–
	[6.769]		
Lagged Volatility	–	0.627	–
		[7.747]	
Lagged Third Moment$^{1/3}$	–	–	0.229
			[2.816]

Panel B	Mean	Volatility	Third Moment$^{1/3}$
Lagged Mean	0.592	0.001	−0.153
	[8.432]	[0.023]	[−1.564]
Lagged Volatility	0.186	0.643	−0.106
	[0.565]	[7.314]	[−0.536]
Lagged Third Moment$^{1/3}$	0.489	0.035	0.162
	[2.306]	[0.723]	[1.799]

Panel C	Mean	Volatility	Third Moment$^{1/3}$
Mean	1.000	−0.288	−0.190
Volatility		1.000	−0.238
Third Moment$^{1/3}$			1.000

Notes - Time series properties of cross-sectional moments. Panel A reports the estimates of the AR(1) coefficients for the mean, volatility, and third centered moment to the power of 1/3. Panel B reports the estimates of time series regressions of each variable on the corresponding column and the three lagged variables reported in the rows. The numbers in brackets underneath each estimate are t-statistics. All standard errors are adjusted for heteroskedasticity.

Panel B of Table 1 further investigates the dynamics by including the lags of all three cross-sectional moments as right hand side variables of the regressions. The interesting finding is that the third moment seems to have predictive power for the conditional mean. More specifically our estimates indicate that following periods of positive asymmetry, the conditional mean increases. This property of the conditional skewness will prove itself important in our theoretical analysis, as news to the shape of the distribution of forecasts will matter in forming the entire stream of future growth prospects. Panel C concludes our preliminary analysis, by documenting that the contemporaneous correlation of the three moments is usually low and slightly negative.

3 The economy

3.1 Preferences

A representative consumer orders consumption profiles according to the following preferences:

$$U_t = (1-\delta)\log C_t + \frac{\delta}{1-\gamma}\log E_t\left[\exp\left\{(1-\gamma)U_{t+1}\right\}\right] \tag{1}$$

where γ indexes risk aversion toward atemporal gambles. This specification is due to Hansen and Sargent (1995) and it is the special case of Epstein and Zin (1989) preferences of unit intertemporal elasticity of substitution. These preferences are also known as risk sensitive preferences and have been employed by Anderson (2005) and Tallarini (2000) among others. The key feature of these preferences is that they allow agents to be risk averse in future utility in addition to future consumption. For the purpose of this article, it is convenient to look at a third order Taylor expansion of (1) about the conditional expectation of U_{t+1}:

$$U_t \approx (1-\delta)\log C_t + \delta E_t\left[U_{t+1}\right] + \frac{\delta(1-\gamma)}{2}V_t\left[U_{t+1}\right] + \frac{\delta(1-\gamma)^2}{3}E_t\left[\left(U_{t+1}-E_t\left[U_{t+1}\right]\right)^3\right].$$

This approximation highlights several important aspects of our specification. When $\gamma = 1$, the standard case of time-additive preferences attains: agents care only about high expected utility levels. However, for levels of risk aversion in excess of 1, our consumers care also about smooth future utility (low $V_t\left[U_{t+1}\right]$), and they dislike negative skewness of their future utility profiles.

The variance and negative skewness aversion are specific to this type of preferences and they suggest that skewness of future growth prospects may matter through a variety of channels in our economy. In this paper, we focus on a class of models in which consumption growth is predictable. In particular, we analyze a specification for the dynamics of the predictable component of consumption growth rates in which news about expected future growth prospects are drawn from a skewed distribution, a prediction that conforms well with the empirical evidence reported in section 2.

3.2 Endowments

In this model skewness is directly built into the dynamics of expected consumption growth. As investors are looking at the one-period ahead distribution of macroeconomic growth prospects, they will act under the assumption that such distribution is not normal. Specifically, we show that a parsimonious way of incorporating this idea within a consumption based asset pricing model consists in using a skew-normal distribution with time-varying parameters.

Preliminaries and notation. At each date t, the representative consumer observes a cross-section of n one period ahead consumption growth forecasts: $\{E_t^i(\Delta c_{t+1})\}_{i=1}^n$.[3] We assume that each forecast is a noisy signal of the actual expected one period ahead consumption growth:

$$E_t^i(\Delta c_{t+1}) = E_t(\Delta c_{t+1}) + \xi_t^i, \quad \forall i = \{1, ..., n\}$$

[3]The idea to exploit individual analyst macroeconomic forecasts to infer agent's expectations in asset pricing models is not new. For example, Anderson, Ghysels and Juergens (2009) study asset pricing when agents face risk and uncertainty. They use the degree of disagreement among professional forecasters as a proxy for the amount of uncertainty that agents have about consumption growth in the economy. Via this measure, they empirically demonstrate that uncertainty has a substantial effect on asset prices and find stronger empirical evidence for a uncertainty-return trade-off than for the traditional risk-return trade-off. In this paper, we focus on different features of the cross-section of point forecasts - namely we focus on the skewness of the distribution and its time variation. The distinction is important, as skewness of the cross-section reveals pattern very different from the dispersion of analyst forecasts. The former is a measure of asymmetry, whereas dispersion measures cross-sectional variances.

where the distribution of ξ_t^i may be non-Gaussian to reflect the degree of asymmetry of the distribution of expected growth rates. We shall denote the cross sectional moments at each date as:

$$\begin{aligned}
\widehat{E}_t^{cs}\left(\Delta c_{t+1}\right) &= \frac{1}{n}\sum_{i=1}^{n} E_t^i\left(\Delta c_{t+1}\right) \\
\widehat{V}_t^{cs}\left(\Delta c_{t+1}\right) &= \frac{1}{n}\sum_{i=1}^{n}\left[E_t^i\left(\Delta c_{t+1}\right)-\widehat{E}_t^{cs}\left(\Delta c_{t+1}\right)\right]^2 \\
\widehat{S}_t^{cs}\left(\Delta c_{t+1}\right) &= \frac{\frac{1}{n}\sum_{i=1}^{n}\left[E_t^i\left(\Delta c_{t+1}\right)-\widehat{E}_t^{cs}\left(\Delta c_{t+1}\right)\right]^3}{\left(\widehat{V}_t^{cs}\left(\Delta c_{t+1}\right)\right)^{3/2}}
\end{aligned}$$

Consumption. The logarithm of consumption growth evolves according to the following process:

$$\Delta c_{t+1} = E_t\left[\Delta c_{t+1}\right] + \sqrt{\sigma_t^c}\varepsilon_{t+1}^c \tag{2}$$

where ε_{t+1}^c is $i.i.d.$ distributed as a standard normal and the conditional variance σ_{t+1}^c follows an AR(1) process:

$$\sigma_{t+1}^c = (1-\rho_\sigma)\overline{\sigma}_c + \rho_\sigma\sigma_t^c + \sqrt{\sigma_\sigma}\xi_{t+1}^\sigma. \tag{3}$$

At each date the investor must forecast consumption growth for the following period, $E_t\left[\Delta c_{t+1}\right]$. We assume that this task is carried out by using the cross-sectional mean of one period ahead forecasts, $\widehat{E}_t^{cs}\left(\Delta c_{t+1}\right)$. In order to evaluate securities, she also needs to figure out the way in which average consumption growth is going to evolve over time. In the process of coming up with a sequence of one period ahead consumption growth forecasts, she recognizes that at each point in time there is an entire cross-section of one period ahead forecasts. In times in which there is a large dispersion about the average growth rate ($\widehat{V}_t^{cs}\left(\Delta c_{t+1}\right)$ is large), she forecasts that the uncertainty about the average forecast is also going to be large in the future. Similarly, at times of heightened asymmetry about the consensus prediction ($\widehat{S}_t^{cs}\left(\Delta c_{t+1}\right)$ is large), the agents thinks that it is more likely that the future average forecast will take on extreme values in one direction or the other.

One way to formalize this economic idea is to specify the expected growth process as one in which innovations follow a skew-normal distribution with time-varying parameters. Specifically, define $E_t\left[\Delta c_{t+1}\right] = \mu_c + x_t$ and let

$$x_{t+1} \quad = \quad \rho_x x_t + \varphi_e \sqrt{\sigma_t^x} \varepsilon_{t+1}^x \tag{4}$$

where the innovations ε_{t+1}^x are orthogonal to the innovations to the consumption process, ε_{t+1}^c, and have the following conditional distribution:

$$\varepsilon_{t+1}^x | I_t \sim SKN(0, 1, \nu_{t+1}) \quad .$$

The notation $SKN(0, 1, \nu_{t+1})$ stands for skew-normal distribution with parameters 0, 1, and ν_{t+1} as defined by Azzalini (1985).[4] The skew-normal distribution provides a convenient way of characterizing departures from normality which may consist in negatively or positively skewed shocks (see Figure 2). The convenience of this distribution also stems from the fact that the first four centered moments are available in closed form. Furthermore, we show in the Appendix that the exponential of the level and the square of a skew-normal random variable can be computed in closed form, which adds to the computational appeal of this distribution.[5] We shall assume that

[4]See also Carmichael and Cóen (2011) who used the skew-normal distribution in reduced form asset pricing models. In contrast, our approach focuses on modeling skewness in expected macro fundamentals.

[5]Specifically: $E_t\left(\varepsilon_{t+1}^x\right) = \sqrt{\frac{2}{\pi}}\, E_t \phi_{t+1}$, $V_t\left(\varepsilon_{t+1}^x\right) = 1 - \frac{2\left(E_t \phi_{t+1}\right)^2}{\pi}$, and $S_t\left(\varepsilon_{t+1}^x\right) = \frac{4-\pi}{2} \frac{\left(\sqrt{2/\pi}\, E_t \phi_{t+1}\right)^3}{\left(1 - 2\left(E_t \phi_{t+1}\right)^2/\pi\right)^{3/2}}$, where $S_t\left(\varepsilon_{t+1}^x\right)$ denotes the third conditional standardized moment and $\phi_t = \frac{\nu_t}{\sqrt{1+\nu_t^2}}$. It is straightforward to show that $S_t\left(\varepsilon_{t+1}^x\right)$ is an monotonically increasing function of ν_t, when ν_t is stationary.

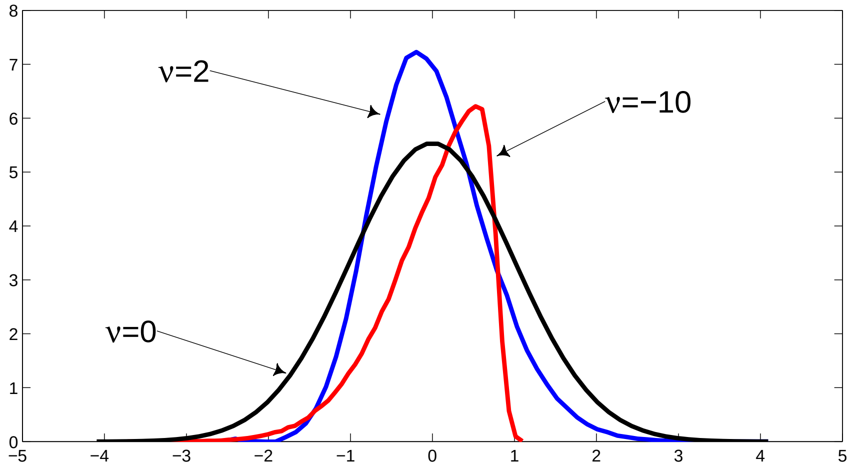

FIG. 2 - The Skew-Normal distribution for different degrees of asymmetry. The parameter ν governs the skewness of the distribution.

ν_t, which governs the degree of skewness, follows an AR(1) process

$$\nu_{t+1} \quad = \quad \rho_\nu \nu_t + \sqrt{\sigma_\nu} \xi_{t+1}^{\nu} \tag{5}$$

To reduce the dimensionality of the model, we rescale the process for σ_t^x in such a way that the innovations to consumption growth and to consumption growth forecasts are proportional to each other:

$$\sigma_t^x = \sigma_t^c / \left(1 - \frac{2 \left(E_t \phi_{t+1} \right)^2}{\pi} \right)$$

where $\phi_t = \frac{\nu_t}{\sqrt{1+\nu_t^2}}$. Given this specification of the model, the first three moments of the distribution of consumption growth forecasts are time-varying. We show in the

Appendix that:

$$
\begin{aligned}
E_t(x_{t+1}) &= \rho_x x_t + \left(\frac{2}{4-\pi}\right)^{1/3} V_t(x_{t+1})^{1/2} |S_t(x_{t+1})|^{1/3} \mathbf{sign}(S_t(x_{t+1})), \qquad (6)\\
V_t(x_{t+1}) &= \varphi_e^2 \sigma_t^c,\\
S_t\left(x_{t+1}\right) &= \frac{4-\pi}{2} \frac{\left(\sqrt{2/\pi}\, E_t \phi_{t+1}\right)^3}{\left(1 - 2\left(E_t \phi_{t+1}\right)^2/\pi\right)^{3/2}}.
\end{aligned}
$$

4 A calibrated economy

Baseline calibration. Table 2 reports our baseline calibration. The model is calibrated to describe a monthly decision problem. We approximate σ_t^c and ν_t on discrete grids and assume independent Markov transition processes for their dynamics. Specifically, we adopt the Rouwenhorst method to approximate AR(1) transition dynamics with various degrees of persistence. Kopecky and Suen (2010) describe this procedure. We approximate the σ_t^c process by a symmetric and evenly-spaced state space $Y_N = \{y_1, \ldots, y_N\}$, with $N = 21$ defined over the interval $[0, 2\psi]$. The transition matrix Θ_N with two parameters $p,\ q \in (0,1)$ is defined recursively as follows:

$$
p\begin{bmatrix} \Theta_{N-1} & \mathbf{0} \\ \mathbf{0}' & 0 \end{bmatrix} + (1-p)\begin{bmatrix} \mathbf{0} & \Theta_{N-1} \\ 0 & \mathbf{0}' \end{bmatrix} + q\begin{bmatrix} \mathbf{0}' & 0 \\ \Theta_{N-1} & \mathbf{0} \end{bmatrix} + (1-q)\begin{bmatrix} 0 & \mathbf{0}' \\ \mathbf{0} & \Theta_{N-1} \end{bmatrix},
$$

where

$$
\Theta_2 = \begin{bmatrix} p & 1-p \\ 1-q & q \end{bmatrix}
$$

and $\mathbf{0}$ is an $(N-1)$-by-1 vector of zeros. Kopecky and Suen (2010) show that using the persistence ρ_σ and shock volatility σ_σ alone, it is possible to construct the approximate Markov chain, with $p = q = (1+\rho_\sigma)/2$ and $\psi = \sqrt{(N-1)\sigma_\sigma/(1-\rho_\sigma^2)}$. Using the semi-annual frequency Livingston dataset, we calibrate the parameters of the model at a monthly frequency. In particular, we set $\rho_\sigma = 0.65^{1/6} = 0.93$, and $\sigma_\sigma^{1/2} = 3.80 \times 10^{-6}$.

TABLE 2
BASELINE CALIBRATION

γ	Risk aversion	10
δ	Subjective discount factor	0.998
μ_c	Average consumption growth	0.001
ρ_x	Autoregressive coefficient of the expected consumption growth rate x_t	0.9619
ϕ_e	Ratio of long-run shock and short-run shock volatilities	0.05
μ_x	Location parameter of skew normal distribution of the innovations to x_t	0
$\sqrt{\sigma_\sigma}$	Conditional volatility of the variance of the short-run shock to consumption growth	3.80×10^{-6}
ρ_σ	Persistence of the variance of the short-run shock to consumption growth	0.93
$\sqrt{\sigma_\nu}$	Conditional volatility of the scale parameter ν of the skew normally distributed innovations to x_t	0.4696
ρ_ν	Persistence of the scale parameter ν of skew normally distributed innovations to x_t	0.8
λ	Leverage	3

Notes - the calibration is set to describe a monthly decision problem.

As a consequence, the 21-state discrete Markov process has parameters $p = q = 0.965$ and $\psi = 4.6235 \times 10^{-5}$.

Following the same procedure, we approximate the process of the variable that governs the skewness dynamics, ν_t, with a symmetric state space $Z_N = \{z_1, \ldots, z_N\}$ with $N = 21$ evenly spaced nodes over the interval $[-\phi, \phi]$. Using the Livingston dataset, we estimated the persistence and unconditional volatility of the skewness process $S_t(x_{t+1})$, which is an increasing and nonlinear function of ν_t. For all the calibrations in this paper, we find that $S_t(x_t)$ can accurately be approximated as an AR(1) process, provided that ν_t is also an AR(1) process. Specifically, we set the monthly persistence of ν_t to 0.8, and its volatility to 0.4696, by using a 21-state discrete Markov process has parameters $p = q = (1 + \rho_\nu)/2 = 0.9$ and $\phi = \sqrt{20}\sqrt{\sigma_\nu/(1 - \rho_\nu^2)} = 3.5$. This calibration results in an AR(1) process for skewness, which is line with our estimates in section 2. In the benchmark model, we assume that skewness is on average zero, but we explore the case of average negative skewness in the section that describes the sensitivity analysis.

The calibration of the other parameters is standard in the long-run risks literature. In particular, we set the persistence of the predictive component of consumption growth, ρ_x, to 0.9612. This value is within the significance range given the estimates that we provided in section 2, and it is overall on the low end of the typical values which are typically find in this literature.

Consumption. Table 3 reports several moments of the distribution of consumption growth and its conditional mean for various horizons. This exercise is relevant, because we need to make sure that the time-variation in the first three moments of the conditional mean of consumption that we have parameterized in the previous sections produces consumption dynamics which are consistent with the observed data.

The dynamics of the the first three moments of the distribution of the conditional mean produce a process of consumption growth which is consistent with the observed dynamics of annual US historical data. It is worth commenting on the negative skewness and excess kurtosis that we find in the data. These are due mostly to the fact that we focus on the largest possible sample of US data, which starts in 1929. The inclusion of the Great Depression on the 1930s' and the World War II years are responsible for the reported estimation of third and fourth moments. We decided not to pursue negative skewness and excess kurtosis in the treatment of our model for practical computational purposes, but we would expect our result to be even more dramatic with the inclusion of these unconditional non-normalities.

Also note that the inclusions of the three persistent components in the dynamics of consumption growth does not produce excessive autocorrelation in consumption growth. The model is simulated at a monthly frequency and aggregated to annual frequency: the degree of persistence of consumption is very much in line with US historical data.

TABLE 3
TIME SERIES PROPERTIES OF CONSUMPTION GROWTH

	Data		Model		
	Estimate	S.E.	Mean	2.5%	97.5%
$E[\Delta c]$	2.042	(0.259)	1.189	0.096	2.282
$\sigma[\Delta c]$	2.285	(0.366)	2.869	2.307	3.431
$skew[\Delta c]$	−1.629	(0.277)	0.012	−0.519	0.543
$kurt[\Delta c]$	7.971	(0.555)	2.933	2.004	3.862
$AC_1[\Delta c]$	0.505	(0.085)	0.569	0.354	0.783
$AC_5[\Delta c]$	−0.061	(0.117)	0.009	−0.241	0.260
$AC_{10}[\Delta c]$	0.006	(0.121)	−0.025	−0.233	0.183

Notes - The table reports the unconditional mean, volatility, skewness, and kurtosis for real US log-consumption growth computed using annual data from 1929 to 2006. The column labeled "S.E." reports the standard errors of these moments. The columns labeled "Mean", "2.5%", and "97.5%" report the mean, bottom 2.5%, and top 97.5% of the distribution of the corresponding moment, obtained from simulating the model 1000 times with sample size 100 years. Consumption is temporally aggregated to annual frequency.

Equilibrium Utility. We shall solve for the utility minus log-consumption:

$$V_t = U_t - \log C_t = \delta\theta \log E_t \exp\left\{\frac{V_{t+1} + \Delta c_{t+1}}{\theta}\right\}$$

We document in the Appendix that V_t can be decomposed as the sum of two terms. The first one is linear in x_t, while the other one is non-linear in σ_t and ϕ_t:

$$V_t = \frac{\delta}{1 - \delta\rho_x} \cdot x_t + \widetilde{V}(\sigma_t, \nu_t).$$

We will use the notation $\widetilde{V}(\sigma_t, \nu_t)$ and $\widetilde{V}_t$ interchangeably. Figure 3(a) shows $\widetilde{V}_t$ as function of σ_t and ν_t. The value function can take on a large range of values as the conditional skewness and volatility explore the state space. Figure 3(b) is even more insightful. This panel reports three horizontal cuts of the utility function. The middle line refers to the case of zero skewness. This is a version of the Bansal and Yaron (2004) model. Notice that in this case the value function is not extremely sensitive to changes in volatility, given our chosen calibration, that postulates a very

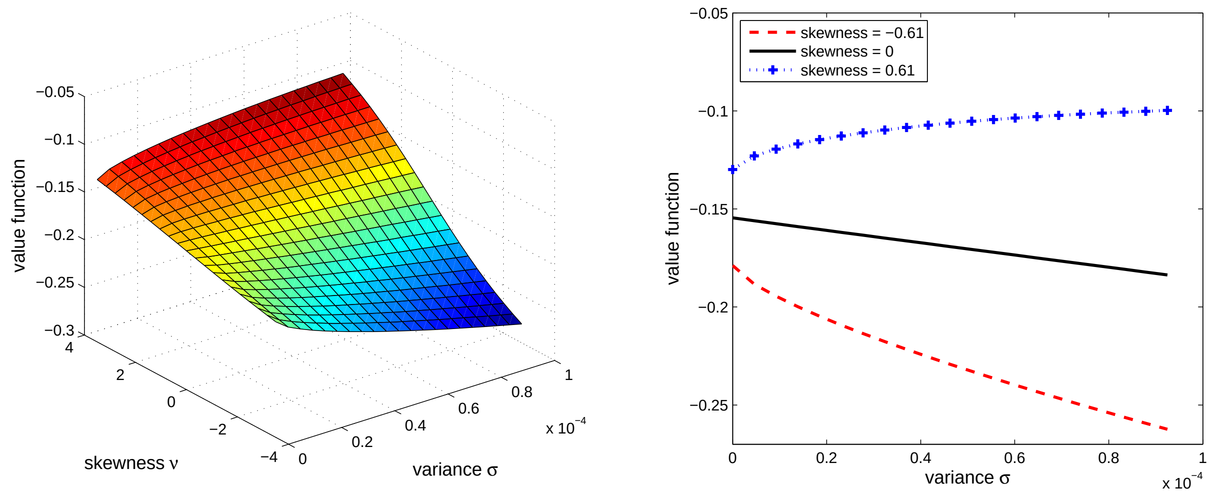

FIG. 3 - Value function. The left panel reports $\widetilde{V}_t$ as a function of the skewness and of the variance parameters. The right panel shows three slices of $\widetilde{V}_t$ for different values of skewness.

small amount of time-varying volatility. The situation is very different for the cases in which skewness is positive or negative. The interaction between second and third moments produces large movements in total discounted utilities. As the degree of asymmetry gets more and more positive, volatility is welfare increasing as it implies a larger probability of landing in an extremely good state of the economy. The opposite is true for the case of negative skewness, since more volatility increases the likelihood of a left tail event. The important message is that time-variation in skewness amplifies the magnitude of utility fluctuations. We shall see in the next section how important this is for the market price of risk.

5 Asset Pricing

We divide our analysis of the asset pricing properties of the model in two parts. First, we study the properties of the stochastic discount factor through which future uncertain payoffs are being discounted by time and by risk. In particular, we check the ability of the model to satisfy the Hansen and Jagannathan (1991) volatility bound and the entropy bound recently proposed by Backus, Chernov, and Zin (2012). Second, we study the properties of the returns of a claim to levered equity and document that it is possible to quantitatively replicate several properties of the distribution of equity returns.

5.1 Bounds

Hansen-Jagannathan bound. Hansen and Jagannathan (1991) construct bounds on the first and second moments of stochastic discount factors that are consistent with the a given distribution of payoffs on a set of primitive securities. Let $\boldsymbol{R}$ denote the vector of quarterly returns of the S&P500 index and the 3 months Treasury bill and let $E[\boldsymbol{R}]$ and $cov(\boldsymbol{R}, \boldsymbol{R})$ be the vector of expected returns and covariance matrix, respectively. Then the lower bound for the volatility of the stochastic discount factor is

$$\sigma(M) \geq \sqrt{(\mathbf{1} - E[M]\, E[\boldsymbol{R}])'\, cov(\boldsymbol{R}, \boldsymbol{R})^{-1} (\mathbf{1} - E[M]\, E[\boldsymbol{R}])}$$

where $E[M]$ is the expected value of the stochastic discount factor and $\mathbf{1}$ is a vector of ones.

The stochastic discount factor can be calculated as the intertemporal marginal rate of substitution:

$$M_{t+1} \;=\; \frac{\partial U_t / \partial C_{t+1}}{\partial U_t / \partial C_t} = \exp\left\{ \log\delta - \Delta c_{t+1} + \frac{U_{t+1}}{\theta} - \log E_t \exp\left\{ \frac{U_{t+1}}{\theta} \right\} \right\}. \tag{7}$$

The Appendix reports the details of the calculations. Figure 4 shows the lower bound on volatility as a function of the average of the stochastic discount factor along with the pairs obtained for several calibrations of the model and by letting the coefficient of risk aversion γ vary between 1 and 20. The bound was obtained using US quarterly data on the S&P500 index and three months Treasury bills. The steepest line refers to the baseline Bansal and Yaron (2004) model, in which any time variation in volatility has been shut down. Notice that a coefficient of risk aversion of about 15-16 is needed for the model to deliver a pair into the acceptable region. The introduction of time-varying volatility (line with circles) is beneficial, in that the mean of the stochastic discount factor increases (thereby reducing the average risk-free rate) and its volatility also increases somewhat for any given degree of risk aversion. The introduction of time-varying skewness (lines with X's and triangles) produces a dramatic increase in the volatility of the stochastic discount factor. Note that for the two calibrations with skewness, a risk aversion of only 9-10 is now needed to get within the Hansen and Jagannathan acceptance region. For the baseline calibration the increase in volatility is so large, that for given γ equity Sharpe ratios up to 50% larger can be achieved relative to a model without any time-variation in skewness.

Entropy bound. Backus, Chernov, and Zin (2012) define the conditional entropy of the pricing kernel as:

$$L_t\left(M_{t+1}\right) = logE_t M_{t+1} - E_t \log M_{t+1}$$

and show that, together with the Euler equation $E_t\left[M_{t+1}R_{t+1}\right] = 1$, it leads to the entropy bound:

$$EL\left(M_{t+1}\right) \quad \geq \quad E\left(\log R_{t+1} - r_{f,t}\right) \tag{8}$$

where $r_{f,t}$ is the logarithm of the one period risk free rate. Equation 8 has the interpretation that mean excess returns are bounded above by the mean conditional entropy of the pricing kernel.

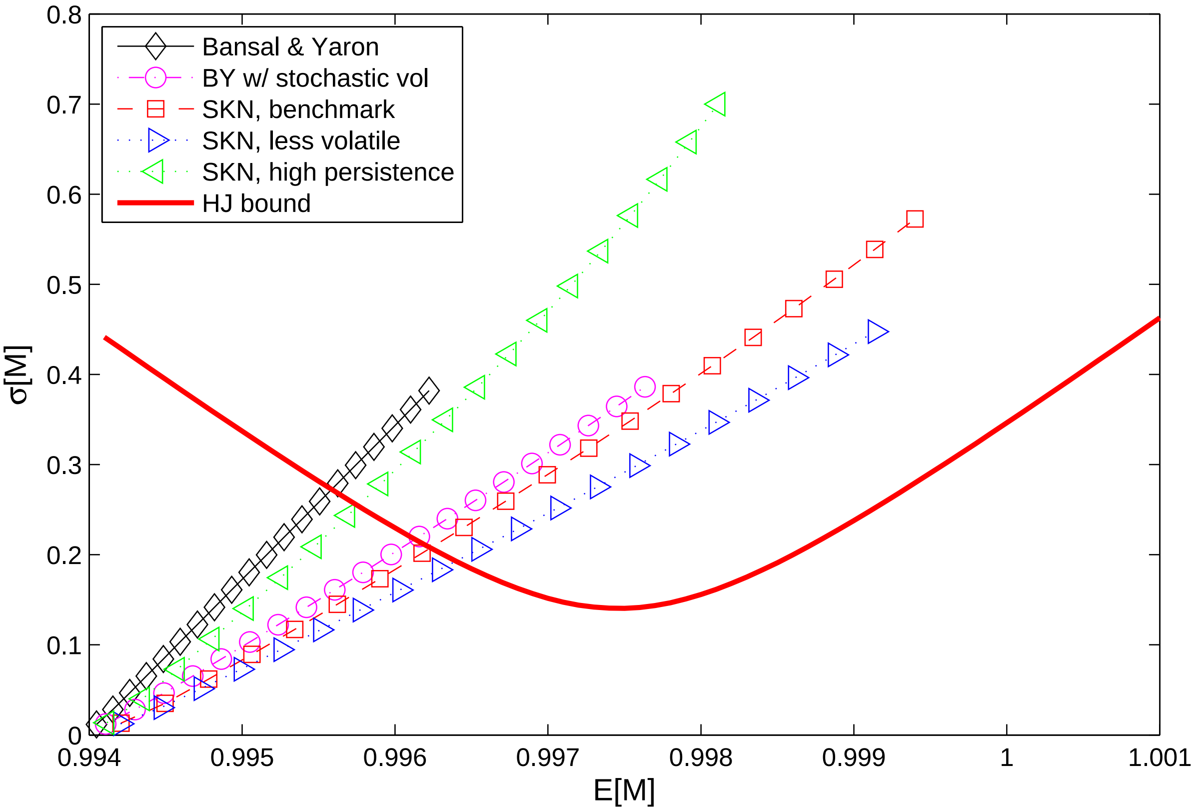

FIG. 4 - Hansen-Jagannathan volatility bound. The thick line is the lower bound calculated using US quarterly data on the S&P500 index and three months Treasury bills. Each line refers to the mean-volatility pairs obtained for the model's calibration reported in the top-left corner. "B&Y model w/constat vol" refers to a model calibrated as in the benchmark, except for the volatility being constant and the skewness being equal to zero at all times. "B&Y model w/stochastic vol" refers to a model calibrated as in the benchmark, except for the skewness being equal to zero at all times. The third model refers to the benchmark calibration, while in the last one the persistence of the ν process is lower and so is its volatility. Each point on the lines refers to an increasing coefficient of risk aversion ($\gamma \in \{1, 20\}$).

We document in the appendix that

$$EL(M_{t+1}) \quad = \quad -\frac{1}{\theta}(\mu_c + \bar{\sigma}) + \frac{1}{2}\bar{\sigma} + \frac{1-\delta}{\delta} E\left[\tilde{V}_{t+1}\right] \tag{9}$$

where $E\left[\tilde{V}_{t+1}\right]$ denotes the unconditional mean of $\tilde{V}_{t+1}$. Figure 5 reports the entropy bound in equation (9) for increasing values of γ and for the same specifications of the model discussed in the previous subsection. The figure confirms the ability of the model to satisfy the bound for degrees of risk aversion as low as 6. Equivalently,

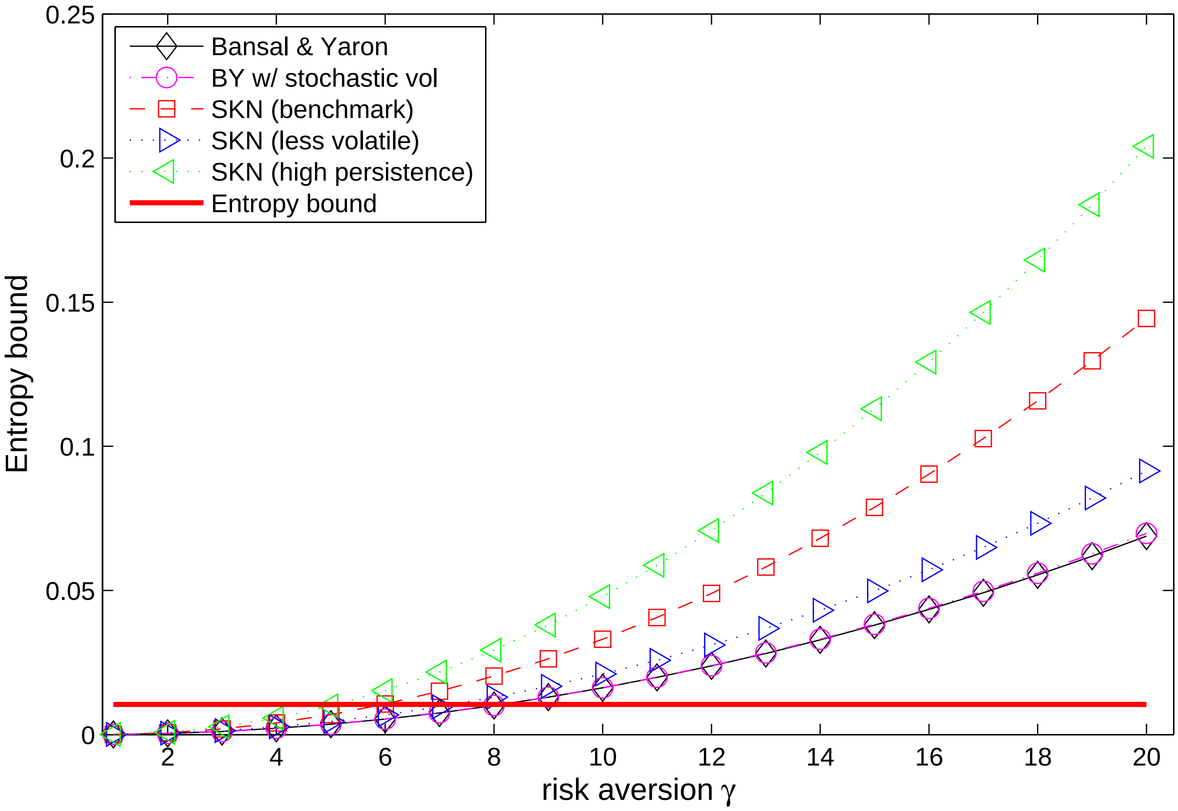

FIG. 5 - Entropy bound. The thick line is the lower bound calculated using US quarterly data on the S&P500 index and three months Treasury bills. Each line refers to the entropy bound for the model's calibration reported in the top-left corner. "B&Y model w/constat vol" refers to a model calibrated as in the benchmark, except for the volatility being constant and the skewness being equal to zero at all times. "B&Y model w/stochastic vol" refers to a model calibrated as in the benchmark, except for the skewness being equal to zero at all times. The third model refers to the benchmark calibration, while in the last one the persistence of the ν process is lower and so is its volatility. Each point on the lines refers to an increasing coefficient of risk aversion ($\gamma \in \{1, 20\}$).

for the same amount of risk aversion needed in a model without time -varying skewness to satisfy the entropy bound, the model presented in this paper can deliver a maximum equity risk premium between 2 and 3 times larger.

5.2 Time series properties of equity returns

We study the properties of the returns to a claim to levered consumption, that is a cash flow whose dynamics are defined as $\Delta d_t = \lambda \Delta c_t$, with $\lambda = 3$. The returns to this claim, R^d, satisfy an Euler equation $E_t \left[M_{t+1} R^d_{t+1} \right] = 1$, where M_{t+1} is the stochastic discount factor reported in the previous section. The details of how to solve the above Euler equation are reported in the Appendix. In Table 4 we report the results of our benchmark calibration in the column labeled "Benchmark". For comparison, we also report the actual moments calculated using annual US data from 1929 to 2006, as well as two alternative calibrations: one in which skewness is made more volatile by increasing the parameter σ_ν (labeled "Volatile Skewness") and one in which any time-variation in skewness in shut down, while keeping the volatility process alive (labeled "No Skewness"). The results are obtained from simulating the models 1000 times with sample size equal to 100 years.

Several results ought to be noticed. First of all, notice that the introduction of skewness determines an increase in the average equity risk premium which doubles the equity premium in the absence of skewness dynamics. This increase comes together with more volatile equity excess returns, with an overall increase in equity Sharpe ratios in the order of 30% to 40%. Second, the average risk free rate is almost unaffected by the introduction of skewness dynamics. Its volatility increases in the two skewness calibrations, but the 95% confidence intervals (reported underneath each estimate) reveal that these increases are well within the margin of significance. Last, the average price-dividend ratio is even closer to data thanks to the introduction of the time-varying skewness process, and so are its volatility and autocorrelation.

Sensitivity analysis. Table 5 documents the sensitivity of our results to several alternative calibrations. Specifically, we consider three main specifications, in which we alter the degree of persistence of the predictive component, ρ_x, and the average

TABLE 4
BENCHMARK CALIBRATION: RESULTS

	Data	Benchmark	Volatile Skewness	No Skewness
$E[r_t^d - r_t^f]$	6.33	7.80	8.83	2.89
		$[5.46, 10.15]$	$[6.39, 11.27]$	$[1.04, 4.74]$
$\sigma[r_t^d - r_t^f]$	19.4	16.0	18.2	9.30
		$[13.8, 18.2]$	$[15.8, 20.6]$	$[7.93, 10.7]$
$E[r_t^f]$	1.16	1.89	1.89	1.89
		$[0.857, 2.93]$	$[0.754, 3.03]$	$[1.26, 2.51]$
$\sigma[r_t^f]$	1.89	2.22	2.44	1.37
		$[1.68, 2.76]$	$[1.85, 3.03]$	$[1.04, 1.70]$
$E[p/d]$	3.30	2.82	2.66	4.47
		$[2.75, 2.88]$	$[2.59, 2.74]$	$[4.43, 4.50]$
$\sigma[p/d]$	0.312	0.169	0.191	0.089
		$[0.137, 0.201]$	$[0.157, 0.226]$	$[0.071, 0.107]$
$AC_1[p/d]$	0.870	0.516	0.502	0.521
		$[0.331, 0.701]$	$[0.316, 0.688]$	$[0.338, 0.704]$

Notes - The first column reports the statistics of interest calculated using annual US data from 1929 to 2006. The second column reports the results from the model using the benchmark calibration. The column labeled "Volatile Skewness" refers to the becnhmark calibration with $\sqrt{\sigma_\nu}$ set to 0.604, instead of 0.469. The column label "No Skewness" refers to the benchmark calibration with $\sqrt{\sigma_\nu}$ and ρ_ν equal to zero. The numbers in squared brackets underneat each statistic are 95% confidence intervals obtains from 1000 simulations of sample size 100 years.

volatility of the shocks, $\sqrt{\sigma_c}$.[6] We label the three cases as "Benchmark" ($\rho_x = 0.962$ and $\sqrt{\sigma_c} = 0.0068$), "Medium Persistence" ($\rho_x = 0.969$ and $\sqrt{\sigma_c} = 0.0058$), and "High Persistence" ($\rho_x = 0.979$ and $\sqrt{\sigma_c} = 0.0058$). For each case, we report the results for a number of possible combinations of the parameters that govern the skewness dynamics (ρ_ν ranging from 0.8 to 0.86, and $\sqrt{\sigma_\nu}$ ranging between 0.2 and 0.6), as well as the calibration in which skewness is fixed at zero.

The main messages looking at the three panels of Table 5 seem to be that Sharpe ratios increase on average by 50% thanks to the introduction of skewness dynamics

[6] We adjust $\sqrt{\sigma_c}$ in such a way that increasing the persistence parameter ρ_x does not alter too much the overall volatility of consumption growth.

and in some cases they can even get three times as large relative to the zero skewness specification. The volatility of consumption growth is usually moderately low, as the 95% confidence intervals from the simulations typically include the number estimated from actual data. For some of the most extreme calibrations Panel C documents that the autocorrelation of consumption growth becomes excessively large, but this is generally not an issue for the "Benchmark" and "Medium Persistence" calibrations.

Empirical predictions. A well established empirical fact in the asset pricing literature is the tendency of equity risk premia and returns' volatilities to vary over time. It is therefore natural to ask whether this model can produce any systematic time-variation in the first two moments of equity excess returns. In Table 7 we report the regressions for equity excess returns and their realized variances on the lagged values of the conditional mean of consumption growth, its conditional variance, and its conditional skewness. Panel A documents that the conditional moments of expected consumption growth can predict future values of equity returns. In particular it seems that the odd moments (mean and skewness) have negative signs and the even moment (variance) has a positive sign in our regressions. The explanation for this is that better growth prospects, in the sense of better average forecasts and increased upside potential, decrease the conditional premium requested for holding risky assets. Similarly, more uncertain growth opportunities determine an increase in conditional equity risk premia, a result already set forward by Bansal and Yaron (2004). We repeat this exercise in Panel B of Table 7 by changing the dependent variable to the realized variance of equity excess returns. The results clearly indicate that the most significant variable in this set of regressions is the conditional variance of expected consumption growth. In the next section we explore the validity of these prediction of the model, by employing the Livingston dataset.

TABLE 5
SENSITIVITY ANALYSIS

PANEL A: SHARPE RATIOS

	Benchmark ρ_ν 0.80	0.82	0.86	*Medium Persistence* ρ_ν 0.80	0.82	0.86	*High Persistence* ρ_ν 0.80	0.82	0.86
$\sqrt{\sigma_\nu}$ 0.20	41.00	42.19	45.50	40.30	41.60	45.24	54.84	56.86	62.57
$\sqrt{\sigma_\nu}$ 0.47	51.62	54.40	62.27	51.88	54.86	63.29	72.70	77.19	89.72
$\sqrt{\sigma_\nu}$ 0.60	55.79	59.07	68.15	56.35	59.86	69.54	79.41	14.88	98.80
No Skew		36.00			24.55			32.39	

PANEL B: VOLATILITY OF CONSUMPTION

	Benchmark ρ_ν 0.80	0.82	0.86	*Medium Persistence* ρ_ν 0.80	0.82	0.86	*High Persistence* ρ_ν 0.80	0.82	0.86
$\sqrt{\sigma_\nu}$ 0.20	2.48 [2.03, 2.94]	2.53 [2.06, 3.01]	2.65 [2.13, 3.17]	2.23 [1.80, 2.67]	2.29 [1.83, 2.74]	2.40 [1.90, 2.91]	2.48 [1.92, 3.04]	2.55 [1.96, 3.15]	2.72 [2.06, 3.38]
$\sqrt{\sigma_\nu}$ 0.47	2.88 [2.31, 3.45]	3.00 [2.39, 3.61]	3.28 [2.58, 3.98]	2.64 [2.08, 3.19]	2.76 [2.15, 3.36]	3.04 [2.33, 3.75]	3.02 [2.27, 3.77]	3.18 [2.37, 3.99]	3.57 [2.58, 4.56]
$\sqrt{\sigma_\nu}$ 0.60	3.05 [2.44, 3.67]	3.19 [2.51, 3.86]	3.50 [2.75, 4.26]	2.81 [2.20, 3.42]	2.94 [2.27, 3.62]	3.26 [2.51, 4.02]	3.25 [2.41, 4.08]	2.87 [2.30, 3.44]	3.85 [2.81, 4.90]
No Skew		2.32 [1.91, 2.73]			1.47 [1.21, 1.74]			1.62 [1.25, 1.98]	

PANEL C: AUTOCORRELATION OF CONSUMPTION

	Benchmark ρ_ν 0.80	0.82	0.86	*Medium Persistence* ρ_ν 0.80	0.82	0.86	*High Persistence* ρ_ν 0.80	0.82	0.86
$\sqrt{\sigma_\nu}$ 0.20	0.45 [0.28, 0.63]	0.47 [0.28, 0.65]	0.49 [0.32, 0.67]	0.50 [0.32, 0.68]	0.51 [0.33, 0.70]	0.54 [0.37, 0.72]	0.58 [0.40, 0.76]	0.60 [0.41, 0.78]	0.63 [0.46, 0.80]
$\sqrt{\sigma_\nu}$ 0.47	0.54 [0.38, 0.71]	0.56 [0.39, 0.73]	0.61 [0.45, 0.78]	0.59 [0.43, 0.76]	0.61 [0.45, 0.78]	0.66 [0.51, 0.82]	0.68 [0.52, 0.84]	0.70 [0.54, 0.85]	0.74 [0.60, 0.88]
$\sqrt{\sigma_\nu}$ 0.60	0.57 [0.41, 0.73]	0.59 [0.43, 0.75]	0.64 [0.49, 0.79]	0.62 [0.46, 0.78]	0.64 [0.49, 0.79]	0.69 [0.54, 0.83]	0.70 [0.56, 0.85]	0.54 [0.37, 0.71]	0.76 [0.63, 0.89]
No Skew		0.40 [0.21, 0.59]			0.44 [0.25, 0.63]			0.52 [0.32, 0.73]	

Notes - The three panels report Sharpe ratios, volatility of consumption growth, and its first order autocorrelation for various calibrations of the model. All numbers are annualized. In each subpanel, "Benchmark" refers to case of $\rho_x = 0.9619$ and $\sqrt{\sigma^c} = 0.0068$, "Medium Persistence" to the case of $\rho_x = 0.969$ and $\sqrt{\sigma^c} = 0.0058$, "High Persistence" to the case of $\rho_x = 0.979$ and $\sqrt{\sigma^c} = 0.0058$. The numbers in brackets are the 95% confidence intervals obtained from 1000 simulations of sample size 100 years.

TABLE 6
MODEL IMPLIED PREDICTIVE REGRESSIONS

	Panel A: Excess Equity Returns		Panel B: Volatility of Equity Returns	
	Coefficients	t-statistic	Coefficients	t-statistic
Intercept	0.0585	4.064	0.3006	20.859
$\widehat{E}_t^{cs}$	-0.0513	-15.388	0.0207	6.204
$\widehat{V}_t^{cs}$	0.0099	3.136	0.0692	21.91
$\widehat{S}_t^{cs}$	-0.0673	-20.172	0.0302	9.052
Adj. R^2	0.1%		1.63%	

Notes - The table reports the model implied predictive regressions for equity excess returns and their realized volatility. All variables are standardized by subtracting their unconditional means and dividing by their standard deviations. The results were obtained by simulating the model at a monthly frequency.

6 Empirical Analysis

We explore the predictive ability of the first three cross-sectional moments of GDP growth forecasts for equity excess returns. We build equity excess returns as the logarithmic difference of the returns on the S&P500 index and the returns on three months Treasury bills. Equity prices are obtained from Shiller's web site, while Treasuries are obtained from the web site of the Federal Reserve Bank of St. Louis. The details and properties of the cross-sectional moments of the distribution of expected real GDP growth rates have been discussed at length in section 2.

Predictive regressions. Table 8 reports the results of our predictive regressions. In all the specifications, we regressed the ex-post six months excess returns on the ex ante cross-sectional moments of the distribution of real GDP growth, and on some additional variables that are known to have predictive power for equity returns. Part of our results confirm the findings of Campbell and Diebold (2009), in that positive average expected GDP growth rates significantly forecast lower future returns, while the opposite is true for the measure of dispersion of forecasts. Furthermore, while the coefficient on average expected growth is always strongly statistically significant, the

TABLE 7
MODEL IMPLIED PREDICTIVE REGRESSIONS

	Panel A: Excess Equity Returns		Panel B: Volatility of Equity Returns	
	Coefficients	t-statistic	Coefficients	t-statistic
Intercept	0.0585	4.064	0.3006	20.859
$\widehat{E}_t^{cs}$	-0.0513	-15.388	0.0207	6.204
$\widehat{V}_t^{cs}$	0.0099	3.136	0.0692	21.91
$\widehat{S}_t^{cs}$	-0.0673	-20.172	0.0302	9.052
Adj. R^2	0.1%		1.63%	

Notes - The table reports the model implied predictive regressions for equity excess returns and their realized volatility. All variables are standardized by subtracting their unconditional means and dividing by their standard deviations. The results were obtained by simulating the model at a monthly frequency.

one on the dispersion is typically not.

The new finding of Table 8 is that skewness also has predictive power for future equity returns. The negative sign of the regression coefficient is also intuitive: a more negative asymmetry suggests an increase in tail risk, and equity holders require extra compensation for it. The coefficient is usually significant at conventional levels and this finding is robust to the inclusion of additional control variables, such as Lettau and Ludvigson (2001) cay, default premium, price-dividend ratio, and term spread.

Taken together, these findings seem to suggest that the odd moments of the distribution of GDP growth forecasts matter in predicting future equity returns, which is consistent with the results of the calibrated model presented in the previous sections.

Table 9 repeats the same exercise for the ex-post realized variance of equity excess returns. Here the situation is reversed, with the dispersion of GDP growth forecast showing up as the only variable with predictive power for future realized variance, a result that seems to be robust to the inclusion of lagged returns' realized variance. Disagreement about future macroeconomic growth prospects is therefore a good indicator of future stock market uncertainty, while the odd moments of the distribution

TABLE 8
PREDICTIVE REGRESSIONS (LEVELS)

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]
Intercept	0.065	0.069	0.068	0.061	0.065	0.067	0.07	0.067	0.068	0.072	0.069	0.07	0.073
	[0.074]	[0.079]	[0.078]	[0.073]	[0.084]	[0.075]	[0.086]	[0.074]	[0.073]	[0.079]	[0.073]	[0.079]	[0.070]
E[growth]	-0.182	-	-	-	-	-	-	-0.17	-0.188	-	-0.178	-	-0.172
	[0.079]							[0.083]	[0.080]		[0.085]		[0.086]
V[growth]	-	0.093	-	-	-	-	-	0.044	-	0.09	0.039	-	0.034
		[0.085]						[0.080]		[0.085]	[0.081]		[0.091]
S[growth]	-	-	-0.104	-	-	-	-	-	-0.115	-0.102	-0.114	-	-0.115
			[0.062]						[0.061]	[0.061]	[0.060]		[0.058]
cay	-	-	-	0.154	-	-	-	-	-	-	-	0.114	0.094
				[0.063]								[0.093]	[0.088]
default	-	-	-	-	0.087	-	-	-	-	-	-	0.022	-0.008
					[0.076]							[0.083]	[0.069]
term pr.	-	-	-	-	-	0.136	-	-	-	-	-	0.155	0.193
						[0.067]						[0.106]	[0.097]
DP	-	-	-	-	-	-	0.145	-	-	-	-	0.184	0.136
							[0.105]					[0.126]	[0.129]
Adj. R^2	0.032	-0.003	0.003	0.017	-0.001	0.014	0.011	0.025	0.043	0.002	0.035	0.046	0.078

Notes - Predictive regressions. For each column the depend variable is the ex-post six months return on the S&P500 index in excess of the risk free rate. $E[growth]$, $V[growth]$, and $S[growth]$ refer to the median, volatility, and skewness of the cross-sectional distribution of expected GDP growth rate at the beginning of each six months interval. The other controls are Leattau and Ludvigson's cay, the term premium, the dividend yield, and the default spread. All standard errors are adjusted for heteroskedasticty.

of expected GDP growth appear not to be playing any significant role in this context.

Skewness and business cycles? The endowment economy that we discussed in the earlier sections allows for a quantitative assessment of the effect of skewness on the conditional and unconditional distribution of equity returns, but it is silent about the economic rationale underlying the specific timing of positive and negative values of the cross-sectional skewness. One possible explanation for why negative skewness commands a positive risk premium is the following. Before the beginning of a recession, the distribution of average forecasts becomes more negatively skewed, as forecasters are expecting their future revisions to become more pessimistic (which tends to be the case during recessions). Equivalently the premium for negative skewness that we observe in the data is a compensation for recession risk. A similar argument can be used to explain why positive skewness reduces the conditional equity risk premium.

We explore this intuition in Figure 6. We construct two dummies for the beginning and for the end of US recessions using the "NBER Business Cycle Expansions and Contractions" dates. To account for the frequency mismatch between the NBER recession dates and the cross-sectional moments of average forecasts (semi-annual), we denote recession semester as a six months span during which the economy was in a recession for at least two months. Also, we omit the January-July 1980 recession from our analysis, because the two dummies for the begging and end of the recession would coincide due to the short duration of the contraction. We then proceed to calculate the correlograms between each of the two recession dummies and the cross-sectional skewness.

Our results seem to indicate the existence of a negative (positive) correlation bewteen skewness and the subsequent start (end) of a recession. This seems to confirm our economic interpretation that skewness becomes more negative before a contraction and an additional equity premium is being requested as a compensation for the

Table 9
Predictive Regressions (Realized Variances)

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Intercept	-0.044	-0.057	-0.044	-0.042	-0.055	-0.041	-0.045	-0.042
	[0.072]	[0.075]	[0.069]	[0.079]	[0.084]	[0.077]	[0.072]	[0.080]
E[growth]	-0.078	-0.105	-	-0.088	-0.128	-	-0.076	-0.087
	[0.059]	[0.061]		[0.065]	[0.073]		[0.060]	[0.065]
V[growth]	0.157	-	0.176	0.182		0.208	0.158	0.182
	[0.078]		[0.076]	[0.090]		[0.092]	[0.079]	[0.089]
S[growth]	-	0.026	0.035	-	0.015	0.027	0.028	0.021
		[0.060]	[0.056]		[0.068]	[0.060]	[0.058]	[0.062]
RV lag	0.085	0.121	0.097	-	-	-	0.088	-
	[0.114]	[0.122]	[0.115]				[0.116]	
Adj. R^2	0.061	0.029	0.052	0.057	0.011	0.044	0.053	0.049

Notes - Predictive regressions. For each column the depend variable is the ex-post realized variance of six months return on the S&P500 index in excess of the risk free rate. $E[growth]$, $V[growth]$, and $S[growth]$ refer to the median, volatility, and skewness of the cross-sectional distribution of expected GDP growth rate at the beginning of each six months interval. "RV lag" is the lagged value of the dependent variable. All standard errors are adjusted for heteroskedasticty.

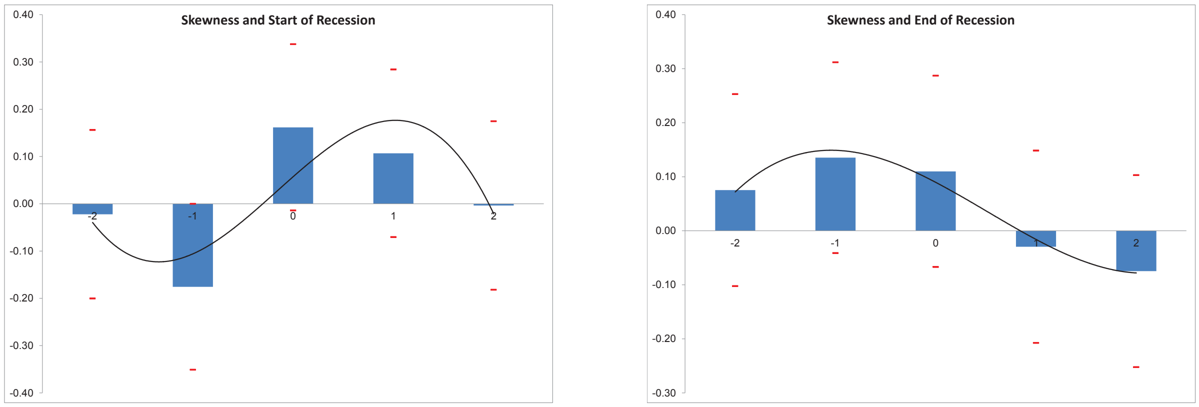

FIG. 6 - Skewness and recessions. In both panels, the bars represent the correlation between a recession dummy and the cross-sectional skewness lag reported on the horizontal axis. In the left panel, the recession dummy indexes the beginning of the recession, while in the right panel it indexes the end of the recession. the horizontal lines above and below each bar represent the 95% confidence interval of the corresponding correlation.

recession that is about to unfold.

7 Concluding remarks

Investors look at the predictions of future growth prospects made by professional forecasters. This paper documents that the entire distribution of such forecasts seems to matter as a larger cross sectional mean, a lower dispersion, and a larger degree of skewness predict lower equity excess returns going forward. The predictive ability of skewness is a novel empirical finding of this paper and it opens up the question of how to think about asymmetric growth prospects in the context of equilibrium asset pricing models. Introducing asymmetry in the distribution of expected consumption growth rates in a way that is consistent with the observed dynamics of consumption produces a sizeable increase in equity Sharpe ratios. Future developments in this literature should look at how these findings generalize to the cross-section of equity returns and to global equity markets.

Appendix

Skew-normal distribution. A skew-normal distribution $SKN(\mu, \sigma, \nu)$ with local parameter μ, scale parameter σ and shape parameter ν, has a probability density function

$$p(x) = \frac{1}{\sigma\pi} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \int_{-\infty}^{\nu\left(\frac{x-\mu}{\sigma}\right)} \exp\left\{-\frac{t^2}{2}\right\} dt.$$

And the first three moments have closed form:

$$\text{mean} = \mu + \sigma\phi\sqrt{\frac{2}{\pi}}, \ \text{variance} = \sigma^2\left(1 - \frac{2\phi^2}{\pi}\right), \ \text{skewness} = \frac{4-\pi}{2}\frac{(\phi\sqrt{2/\pi})^3}{(1-2\phi^2/\pi)^{3/2}},$$

where $\phi = \frac{\nu}{\sqrt{1+\nu^2}}$.

In our model, the innovations ε^x_{t+1} follow the conditional distribution $SKN(0, 1, \nu_{t+1})$. And therefore, the conditional expectation of the shock is:

$$E_t(\varepsilon^x_{t+1}) = \sqrt{\frac{2}{\pi}} E_t\phi_{t+1}.$$

The conditional variance of the shock is:

$$\begin{aligned} V_t(\varepsilon^x_{t+1}) &= E_t[V_t(\varepsilon^x_{t+1}|\nu_{t+1})] + V_t[E_t(\varepsilon^x_{t+1}|\nu_{t+1})] \\ &= E_t\left(1 - \frac{2\phi^2_{t+1}}{\pi}\right) + V_t\left(\sqrt{\frac{2}{\pi}}\phi_{t+1}\right) \\ &= 1 - \frac{2}{\pi}\left(E_t\phi_{t+1}\right)^2. \end{aligned}$$

Similarly, it is straightforward to show that the conditional skewness of the shock is:

$$S_t\left(\varepsilon^x_{t+1}\right) = \frac{4-\pi}{2}\frac{\left(\sqrt{2/\pi}\, E_t\phi_{t+1}\right)^3}{\left(1 - 2\left(E_t\phi_{t+1}\right)^2/\pi\right)^{3/2}}.$$

Solving for the value functions. We shall start by expressing the utility functions in a more convenient form. We define the value functions as the utility minus log-consumption:

$$\begin{aligned} V_t &= U_t - \log C_t \\ &= \delta\theta \log E_t \exp\left\{\frac{V_{t+1} + \Delta c_{t+1}}{\theta}\right\} \end{aligned}$$

We shall decompose V_t as the sum of two terms: one is linear in x_t, the other one is non-linear in σ_t and ν_t:

$$V_t = Bx_t + \widetilde{V}(\sigma_t, \nu_t)$$

we will use the notation $\widetilde{V}(\sigma_t, \nu_t)$ and $\widetilde{V}_t$ interchangeably. It is easy to verify that:

$$B = \frac{\delta}{1 - \delta\rho_x}$$

We shall solve for

$$\widetilde{V}_t \quad = \quad \delta\theta \log E_t \exp\left\{ \frac{\mu_c + \varphi_e B\sqrt{\sigma_t^x}\varepsilon_{t+1}^x + \sqrt{\sigma_t^c}\varepsilon_{t+1}^c + \widetilde{V}_{t+1}}{\theta} \right\} \tag{10}$$

via value function iteration.

In order to solve (10), we are going to use the following lemma.

Lemma 1. *If $z \sim SKN(\mu, \sigma, \nu)$, then*

$$\begin{aligned} \log E \exp\left\{\kappa_1 z + \kappa_2 z^2\right\} \quad &= \quad \log(2) - \frac{1}{2}\log\left(1 - 2\kappa_2\sigma^2\right) + \frac{1}{2\sigma^2}\left[\frac{\left(\kappa_1\sigma^2 + \mu\right)^2}{1 - 2\kappa_2\sigma^2} - \mu^2\right] + \\ &\quad \log\Phi\left(\frac{\nu}{\sigma^2} \cdot \frac{\kappa_1\sigma^2 + \mu - \mu\sqrt{1 - 2\kappa_2\sigma^2}}{\sqrt{1 - 2\kappa_2\sigma^2 + \nu^2}}\right) \end{aligned}$$

for any scalar κ_1 and κ_2, such that $1 - 2\kappa_2\sigma^2 > 0$.

Proof. By definition:

$$\begin{aligned} E\exp\{\kappa_1 z + \kappa_2 z^2\} \quad &= \quad \int_{-\infty}^{+\infty} \exp\{\kappa_1 z + \kappa_2 z^2\} \\ &\quad \left[\frac{1}{\sigma\pi}\exp\left\{-\frac{(z-\mu)^2}{2\sigma^2}\right\} \cdot \left(\int_{-\infty}^{\nu\frac{(z-\mu)}{\sigma}} \exp\left\{-\frac{t^2}{2}\right\} dt\right)\right] dz \end{aligned}$$

where the term in the square brackets corresponds to the probability distribution function of a skew-normal. It follows that:

$$
\begin{aligned}
E\exp\{\kappa_1 z+\kappa_2 z^2\} &= \int_{-\infty}^{+\infty}\frac{1}{\sigma\pi}\exp\left\{-\frac{z^2-2\left(\frac{\kappa_1\sigma^2+\mu}{1-2\kappa_2\sigma^2}\right)z+\frac{\mu^2}{1-2\kappa_2\sigma^2}}{2\left(\frac{\sigma^2}{1-2\kappa_2\sigma^2}\right)}\right\}\cdot\left(\int_{-\infty}^{\nu\frac{(z-\mu)}{\sigma}}\exp\left\{-\frac{t^2}{2}\right\}dt\right)dz \\
&= \exp\left\{\frac{\frac{\left(\kappa_1\sigma^2+\mu\right)^2}{1-2\kappa_2\sigma^2}-\mu^2}{2\sigma^2}\right\}\cdot\int_{-\infty}^{+\infty}\frac{1}{\sigma\pi}\exp\left\{-\frac{\left(z-\left(\frac{\kappa_1\sigma^2+\mu}{1-2\kappa_2\sigma^2}\right)\right)^2}{2\left(\frac{\sigma^2}{1-2\kappa_2\sigma^2}\right)}\right\}\cdot \\
&\quad \left(\int_{-\infty}^{\nu\frac{(z-\mu)}{\sigma}}\exp\left\{-\frac{t^2}{2}\right\}dt\right)dz
\end{aligned}
$$

Apply the following change of variable

$$
y=\frac{z-\left(\frac{\kappa_1\sigma^2+\mu}{1-2\kappa_2\sigma^2}\right)}{\sigma}\cdot\sqrt{1-2\kappa_2\sigma^2}.
$$

Then:

$$
\begin{aligned}
E\exp\{\kappa_1 z+\kappa_2 z^2\} &= \frac{2}{\sqrt{1-2\kappa_2\sigma^2}}\cdot\exp\left\{\frac{\frac{\left(\kappa_1\sigma^2+\mu\right)^2}{1-2\kappa_2\sigma^2}-\mu^2}{2\sigma^2}\right\}\cdot\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{y^2}{2}\right\}\cdot \\
&\quad \left(\int_{-\infty}^{\frac{\nu}{\sqrt{1-2\kappa_2\sigma^2}}y+\frac{(\kappa_1+2\mu\kappa_2)\nu\sigma}{1-2\kappa_2\sigma^2}}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{t^2}{2}\right\}dt\right)dy
\end{aligned}
$$

Now use the following result from Zacks (1981). If U is a $N(0,1)$ random variable, then

$$
E\left[\Phi\left(hU+k\right)\right]=\Phi\left(\frac{k}{\sqrt{1+h^2}}\right)
$$

for any real h,k and where $\Phi(\cdot)$ denotes the cumulative distribution function of a $N(0,1)$ random variable. This implies that:

$$
\begin{aligned}
E\exp\{\kappa_1 z+\kappa_2 z^2\} &= \frac{2}{\sqrt{1-2\kappa_2\sigma^2}}\cdot\exp\left\{\frac{1}{2\sigma^2}\left[\frac{\left(\kappa_1\sigma^2+\mu\right)^2}{1-2\kappa_2\sigma^2}-\mu^2\right]\right\}\cdot \\
&\quad \Phi\left(\frac{(\kappa_1+2\mu\kappa_2)\mu\sigma}{\sqrt{(1-2\kappa_2\sigma^2)(1-2\kappa_2\sigma^2+\nu^2)}}\right)
\end{aligned}
$$

Applying logarithms to both sides concludes the proof. □

For the special case of $\kappa_2=0$ and $\sigma=1$, lemma 1 implies:

$$
\log E_t\exp\{\kappa_1 z\}=\log(2)+\frac{\kappa_1(\kappa_1+2\mu)}{2}+\log\Phi\left(\frac{\nu\kappa_1}{\sqrt{1+\nu^2}}\right)
$$

This allows us to rewrite the right hand side of the Bellman equation as:

$$\begin{aligned}
\widetilde{V}\left(\sigma_t^c, \nu_t\right) &= \delta\theta \log E_t \left\{ E_t \left[\exp \left\{ \frac{\mu_c + \varphi_e B \sqrt{\sigma_t^x} \varepsilon_{t+1}^x + \sqrt{\sigma_t^c} \varepsilon_{t+1}^c + \widetilde{V}_{t+1}}{\theta} \right\} |\nu_{t+1} \right] \right\} \\
&= \delta\mu_c + \frac{\delta}{2\theta}\sigma_t^c + \delta\theta \left(\log(2) + \frac{\varphi_e^2 B^2 \sigma_t^x}{2\theta^2} + \frac{2\mu_x \varphi_e B \sqrt{\sigma_t^x}}{2\theta} \right) + \delta\theta \log E_t \Phi \left(\frac{\nu_{t+1} \varphi_e B \sqrt{\sigma_t^x}}{\theta \sqrt{1 + \nu_{t+1}^2}} \right) \\
&\quad + \delta\theta \log E_t \exp \left\{ \frac{\widetilde{V}\left(\sigma_{t+1}^c, \nu_{t+1}\right)}{\theta} \right\}
\end{aligned}$$

Note that σ_t^x that appears on the right hand side of the previous equation is a function of σ_t^c and ν_t only:

$$\sigma_t^x = \sigma_t^c / \left[1 - \frac{2}{\pi} \left(E_t \frac{\nu_{t+1}}{\sqrt{1 + \nu_{t+1}^2}} \right)^2 \right]$$

Solving for the stochastic discount factors. The stochastic discount factor is the intertemporal marginal rate of substitution:

$$\begin{aligned}
m_{t+1} &= \log \left(\frac{\partial U_t / \partial C_{t+1}}{\partial U_t / \partial C_t} \right) \\
&= \log \delta - \Delta c_{t+1} + \frac{V_{t+1} + \Delta c_{t+1}}{\theta} - \log E_t \exp \left\{ \frac{V_{t+1} + \Delta c_{t+1}}{\theta} \right\} \\
&= \log \delta - \left(1 - \frac{1}{\theta} \right) \Delta c_{t+1} + \frac{V_{t+1}}{\theta} - \frac{V_t}{\delta\theta} \\
&= \log \delta - \left(1 - \frac{1}{\theta} \right) \Delta c_{t+1} + \frac{1}{\theta} \left(\widetilde{V}_{t+1} - \frac{\widetilde{V}_t}{\delta} \right) + \frac{B}{\theta} \left(x_{t+1} - \frac{x_t}{\delta} \right) \qquad (11)
\end{aligned}$$

Notice that the stochastic discount factor captures the innovations in the conditional mean, in the conditional volatility, and in the conditional skewness.

Solving for the risk-free rates. Real risk-free rates can be obtained as:

$$r_t^f = -\log E_t \exp\{m_{t+1}\}$$

Using the stochastic discount factor (11) above:

$$
\begin{aligned}
r_t^f &= -\log(\delta) - \log E_t \exp\left\{\left(\frac{1-\theta}{\theta}\right)\left(\mu_c + x_t + \sqrt{\sigma_t^c}\varepsilon_{t+1}^c\right) + \frac{\left(\widetilde{V}_{t+1} - \frac{\widetilde{V}_t}{\delta}\right)}{\theta} + \frac{B}{\theta}\left(\rho x_t + \varphi_e\sqrt{\sigma_t^x}\varepsilon_{t+1}^x - \frac{x_t}{\delta}\right)\right\} \\
&= -\log(\delta) - (\frac{1}{\theta} - 1)\mu_c + \frac{1}{\theta\delta}\widetilde{V}_t + x_t - \frac{1}{2}\left(\frac{1}{\theta} - 1\right)^2 \sigma_t^c - \log E_t \exp\left\{\frac{B\varphi_e\sqrt{\sigma_t^x}}{\theta}\varepsilon_{t+1}^x + \frac{\widetilde{V}_{t+1}}{\theta}\right\} \\
&= -\log(\delta) - (\frac{1}{\theta} - 1)\mu_c + \frac{1}{\theta\delta}\widetilde{V}_t + x_t - \frac{1}{2}\left(\frac{1}{\theta} - 1\right)^2 \sigma_t^c \\
&\quad - \log E_t\left[E_t\left(\exp\left\{\frac{B\varphi_e\sqrt{\sigma_t^x}}{\theta}\varepsilon_{t+1}^x\right\}|\nu_{t+1}\right)\cdot E_t\left(\exp\left\{\frac{\widetilde{V}_{t+1}}{\theta}\right\}|\nu_{t+1}\right)\right]
\end{aligned}
$$

Rewrite equation (10) and we shall obtain:

$$
\log E_t \exp\left\{\frac{\mu_c + \sqrt{\sigma_t^c}\varepsilon_{t+1}^c + B\varphi_e\sqrt{\sigma_t^x}\varepsilon_{t+1}^x + \widetilde{V}_{t+1} - \widetilde{V}_t/\delta}{\theta}\right\} = 0
$$

Using equation (10) and lemma 1 , we conclude that:

$$
r_t^f = -\log(\delta) + \mu_c + x_t + \left(\frac{1}{\theta} - \frac{1}{2}\right)\sigma_t^c. \qquad \textbf{(12)}
$$

Solving for returns to levered consumption claim. We consider the returns to a claim to levered consumption

$$
\Delta d_t = \lambda \Delta c_t
$$

where $\lambda > 1$ denotes the leverage ratio for the claim on consumption. Then the returns to the levered consumption claim satisfy

$$
\begin{aligned}
1 &= E_t\left[\exp\{m_{t+1}\}\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right)\right] \\
&= E_t\left[\exp\{m_{t+1}\}\left(\frac{1 + \exp\{Vd_{t+1}\}}{\exp\{Vd_t\}}\right)\exp\{\lambda\Delta c_{t+1}\}\right]
\end{aligned}
$$

where $Vd_t = \log(P_t/D_t)$ denotes the log ratio of price to the consumption of levered claims.

Rewrite the above equation, Vd_t is given by:

$$
\begin{aligned}
Vd_t &= \log E_t \exp\left\{m_{t+1} + \lambda \Delta c_{t+1}\right\}\left(1 + \exp\left\{Vd_{t+1}\right\}\right) \\
&= \log\left(E_t \exp\left\{m_{t+1} + \lambda \Delta c_{t+1}\right\} + E_t \exp\left\{m_{t+1} + \lambda \Delta c_{t+1} + Vd_{t+1}\right\}\right)
\end{aligned}
\tag{13}
$$

Using the stochastic discount factor equation (11), the first part in the logarithm of equation (13) is

$$
\begin{aligned}
E_t \exp\left\{m_{t+1} + \lambda \Delta c_{t+1}\right\} &= E_t \exp\left\{\log\delta + \left(\lambda - 1 + \frac{1}{\theta}\right)\left(\mu_c + x_t + \sqrt{\sigma_t^c}\varepsilon_{t+1}^c\right) + \frac{1}{\theta}\left(\widetilde{V}_{t+1} - \frac{\widetilde{V}_t}{\delta}\right)\right. \\
&\quad \left. + \frac{B}{\delta}\left(\rho_x x_t + \varphi_e \sqrt{\sigma_t^x}\varepsilon_{t+1}^x - \frac{x_t}{\delta}\right)\right\} \\
&= 2\delta \exp\left\{\left(\lambda - 1 + \frac{1}{\theta}\right)\mu_c + (\lambda - 1)x_t - \frac{\widetilde{V}_t}{\delta\theta} + \frac{1}{2}\left(\lambda - 1 + \frac{1}{\theta}\right)^2 \sigma_t^c + \frac{B^2\varphi_e^2}{2\theta^2}\sigma_t^x\right\} \cdot \\
&\quad E_t\left[\Phi\left(\frac{B\varphi_e\sqrt{\sigma_t^x}\nu_{t+1}}{\theta\sqrt{1+\nu_{t+1}^2}}\right) \cdot \exp\left\{\frac{\widetilde{V}_{t+1}}{\theta}\right\}\right].
\end{aligned}
$$

We notice that the right hand side of Vd_t expression (13) is a function of state variables $\sigma_t^c,\ \nu_t,\ x_t$. Due to the complexity of the Vd_t function, it is difficult to guess the closed form solution of Vd_t. Therefore, we approximate Vd_t by using a quadratic polynomial function of these three state variables $x_t,\ \sqrt{\sigma_t^c}$, and ν_t:

$$
Vd_t = a_1 + a_2 x_t + a_3 x_t^2 + a_4\sqrt{\sigma_t^c} + a_5\sigma_t^c + a_6\nu_t + a_7\nu_t^2 + a_8 x_t\sqrt{\sigma_t^c} + a_9 x_t\nu_t + a_{10}\sqrt{\sigma_t^c}\nu_t.
$$

Plugging the guess of Vd_t function into the second part in the logarithm of equation (13), we have:

$$
\begin{aligned}
& E_t \exp\left\{m_{t+1} + \lambda \Delta c_{t+1} + Vd_{t+1}\right\} \\
= \ & \delta \exp\left\{\left(\lambda - 1 + \frac{1}{\theta}\right)\mu_c + (\lambda - 1)x_t - \frac{\widetilde{V}_t}{\delta\theta} + \frac{1}{2}\left(\lambda - 1 + \frac{1}{\theta}\right)^2\sigma_t^c\right\} \cdot E_t \exp\left\{\frac{B\varphi_e\sqrt{\sigma_t^x}}{\theta}\varepsilon_{t+1}^x + \frac{\widetilde{V}_{t+1}}{\theta} + Vd_{t+1}\right\} \\
= \ & \delta \exp\left\{\left(\lambda - 1 + \frac{1}{\theta}\right)\mu_c + (\lambda - 1)x_t - \frac{\widetilde{V}_t}{\delta\theta} + \frac{1}{2}\left(\lambda - 1 + \frac{1}{\theta}\right)^2\sigma_t^c\right\} \cdot E_t \exp\left\{\frac{\widetilde{V}_{t+1}}{\theta} + a_1 + a_2 x_t + a_3\rho_x^2 x_t^2\right. \\
& + a_4\sqrt{\sigma_{t+1}^c} + a_5\sigma_{t+1}^c + a_6\nu_{t+1} + a_7\nu_{t+1}^2 + a_8\rho_x x_t\sqrt{\sigma_{t+1}^c} + a_9\rho_x + a_{10}\sqrt{\sigma_{t+1}^c}\nu_{t+1} + a_3\varphi_e^2\sigma_t^x(\varepsilon_{t+1}^x)^2 \\
& \left. + \left(\frac{B\varphi_e\sqrt{\sigma_t^x}}{\theta} + a_2\varphi_e\sqrt{\sigma_t^x}\theta + 2a_3\rho_x\varphi_e\sqrt{\sigma_t^x}\theta x_t + a_8\varphi_e\sqrt{\sigma_t^x}\theta\sqrt{\sigma_{t+1}^c} + a_9\varphi_e\sqrt{\sigma_t^x}\theta\nu_{t+1}\right)\varepsilon_{t+1}^x\right\}.
\end{aligned}
$$

We solve for parameters $a_1,\ a_2,\ a_3,\ a_4,\ a_5,\ a_6,\ a_7,\ a_8,\ a_9,\ a_{10}$, by regressing Vd_t on these three state variables and all their quadratic terms $[1,\ \sqrt{x_t,\ x_t^2,\ \sigma_t^c},\ \sigma_t^c,\ \nu_t,\ \nu_t^2,\ x_t\sqrt{\sigma_t^c},\ \sqrt{\sigma_t^c}\nu_t,\ x_t\nu_t]$, and updating regression coefficients via iteration.

We find that the quadratic polynomial function of $\sqrt{\sigma_t^c},\ \nu_t$, and x_t is a good approximation of Vd_t solution.

The returns to the levered consumption claim

$$r_{d,t+1} = \left(\frac{1+\exp\{Vd_{t+1}\}}{\exp\{Vd_t\}}\right)\exp\{\Delta\lambda c_{t+1}\}$$

can be computed using the Vd_t.

Derivation of Entropy bound. We need to calculate:

$$EL_t(M_{t+1}) = E(logE_tM_{t+1}) - E(\log M_{t+1})$$

We shall focus on the two terms on the right hand side separately. The first one is simply minus the unconditional expectation of the risk-free rate:

$$\begin{aligned} E(logE_tM_{t+1}) &= -E(r_{f,t}) \\ &= \log\delta - \mu_c - \left(\frac{1}{\theta}-\frac{1}{2}\right)\bar{\sigma} \end{aligned} \tag{14}$$

Now turn to the second term:

$$\begin{aligned} E(\log M_{t+1}) &= Em_{t+1} \\ &= \log\delta - \left(1-\frac{1}{\theta}\right)E[\Delta c_{t+1}] + \frac{B}{\theta}Ex_{t+1} - \frac{B}{\delta\theta}Ex_t + \frac{1}{\theta}E\tilde{V}_{t+1} - \frac{1}{\delta\theta}E\tilde{V}_t \\ &= \log\delta - \left(1-\frac{1}{\theta}\right)\mu_c + \frac{1}{\theta}\left(\frac{\delta-1}{\delta}\right)E\left[\tilde{V}_t\right] \end{aligned} \tag{15}$$

where $E\left[\tilde{V}_t\right]$ is the unconditional expectation of $\tilde{V}_t$. This can be calculated by denoting $\bar{\pi}_\sigma$ and $\bar{\pi}_\nu$ as the probability vectors that define the invariant distributions of σ and ν. These are obtained as the eigenvectors (normalized so that the sum of all entries adds up to 1) associated with the unit eigenvalues of the transition matrices Π_σ and Π_ν. By letting the matrix $\tilde{V}(\sigma,\nu)$ be defined as

$$\tilde{V}(\sigma,\nu) = \begin{bmatrix} \tilde{V}(\sigma_1,\nu_1) & \tilde{V}(\sigma_1,\nu_2) & \dots & \tilde{V}(\sigma_1,\nu_N) \\ \tilde{V}(\sigma_2,\nu_1) & \ddots & & \\ \vdots & & & \\ \tilde{V}(\sigma_N,\nu_1) & \dots & & \tilde{V}(\sigma_N,\nu_N) \end{bmatrix}$$

it follows that $E\left[\tilde{V}_t\right] = \bar{\pi}'_\sigma \tilde{V}(\sigma,\nu)\bar{\pi}_\nu$. By combining, (14) and (15), we get

$$EL\left(M_{t+1}\right) = -\frac{1}{\theta}\left(\mu_c + \bar{\sigma}\right) + \frac{1}{2}\bar{\sigma} + \frac{1-\delta}{\delta}E\left[\tilde{V}_{t+1}\right]$$

Calculation of yields. The prices of zero coupon bonds can be computed recursively as

$$\begin{aligned} q_{1,t} &= E_t\left[exp\left\{m_{t+1}\right\}\right] \\ q_{2,t} &= E_t\left[exp\left\{m_{t+1}\right\}q_{1,t+1}\right] \\ &\vdots \\ q_{n,t} &= E_t\left[exp\left\{m_{t+1}\right\}q_{n-1,t+1}\right] \end{aligned}$$

where $q_{j,t}$ is the date t price of a bond with j periods left until maturity. Note that $r_{f,t} = \log\left(1/q_{1,t}\right)$ is the risk-free rate. Log-yields are $y_{j,t} = -\log q_{j,t}$.

References

Anderson, E. (2005). "The dynamics of risk-sensitive allocations," *Journal of Economic Theory*, 125(2), 93–150.

Anderson, E., Ghysels, E. and J. Juergens (2009). "The impact of risk and uncertainty on expected returns," *Journal of Financial Economics*, 94(2), 233–263.

Arditti, F.D., (1967). "Risk and the required return on equity," *Journal of Finance*, 22 (1), 19–36.

Azzalini, A. (1985). "A class of Distributions which Includes the Normal Ones," *Scandinavian Journal of Statistics*, 12, 171–178.

Backus, D., Chernov, M., and S. Zin (2012). "Sources of entropy in representative agent models," *Working Paper*.

Bansal, Ravi and Amir Yaron (2004). "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *The Journal of Finance*, 59, 1481–1509.

Campbell, S. and F. Diebold (2009). "Stock Returns and Expected Business Conditions: Half a Century of Direct Evidence," *Journal of Business and Economic Statistics*, 27, 266–278.

Carmichael, B. and A. Cóen (2011). "Asset Pricing with Skewed Returns," Working Paper - Université de Laval.

Chabi-Yo, F., E. Ghysels, and E. Renault (2010). "Factor asset pricing models implied by heterogeneous beliefs and attitudes towards risk," Working Paper - Ohio State Univeristy and UNC.

Chang, B., P. Christoffersen, and K. Jacobs (2012). "Market Skewness Risk and the Cross-section of Stock Returns," *Journal of Financial Economics* (forthcoming).

Chen, J., H. Hong, H. and J. Stein (2001). "Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices," *Journal of Financial Economics*, 61, 345381.

Conrad, Jennifer, Robert Dittmar, and Eric Ghysels (2009). "Ex ante skewness and expected stock returns," *Journal of Finance* (forthcoming).

Damodaran, A. (1985). "Economic Events, Information Structure, and the Return-Generating Process," *Journal of Financial and Quantitative Analysis* 425434.

Epstein, Larry G. and Stanley E. Zin (1989). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57 (4), 937-69.

Hansen, Lars P. and Ravi Jagannathan (1991). "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.

Hansen, Lars P. and Thomas J. Sargent (1995). "Discounted linear exponential quadratic gaussian control," *IEEE Trans. Automatic Control*, 40(5), 968-971.

Kopecky, K. and R. Suen (2010). "Finite State Markov-chain Approximations to Highly Persistent Processes," *Review of Economic Dynamics*, 13, 701–714.

Kraus, A. and Litzenberger, R.H., (1976). "Skewness preference and the valuation of risk assets," *Journal of Finance*, 31 (4), 1085–1100.

Lettau, M. and Ludvigson, S. (2001). "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, June, 56(3), 815–849.

Mitton, T. and K. Vorkink (2007). "Equilibrium underdiversification and the preference for skewness," *Review of Financial Studies* 20, 1255-1288.

Rubinstein, M. (1973). "The fundamental theorem of parameter-preference security valuation," *Journal of Financial and Quantitative Analysis*, 8 (1), 61–69.

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

Welch, I. and Goyal, A. (2008). "A comprehensive look at the empirical performance of equity premium prediction," *Review of Financial Studies*, 21 (4), 1455–1508.

Zacks, S. (1981). "Parametric statistical inference". Pergamon Press, Oxford.