The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXVIII, NO. 5 • OCTOBER 2013

Uncertainty, Time-Varying Fear, and Asset Prices

ITAMAR DRECHSLER*

ABSTRACT

I construct an equilibrium model that captures salient properties of index option prices, equity returns, variance, and the risk-free rate. A representative investor makes consumption and portfolio choice decisions that are robust to his uncertainty about the true economic model. He pays a large premium for index options because they hedge important model misspecification concerns, particularly concerning jump shocks to cash flow growth and volatility. A calibration shows that empirically consistent fundamentals and reasonable model uncertainty explain option prices and the variance premium. Time variation in uncertainty generates variance premium fluctuations, helping explain their power to predict stock returns.

KNIGHTIAN UNCERTAINTY—AMBIGUITY ABOUT the true model governing fundamentals—can be an important factor influencing investors' consumption and portfolio choice decisions. Incorporating it into asset pricing models can therefore shed light on sources of asset return premiums and time variation in prices. In this paper, I construct an equilibrium model with a representative investor who has model uncertainty that varies in intensity over time.[1] A calibration of this model is shown to capture a broad set of features of equity and equity index option prices. In particular, the model generates two prominent features of index options that represent a challenge to equilibrium models: (i) the large premium embedded in their prices, called the variance premium, and (ii) the shape of their implied volatility surface, especially the "volatility skew." This is the name given to the pattern in the implied volatilities of short-maturity puts, which are high for out-of-the-money (otm) puts and decrease sharply as put moneyness increases. At the same time, the model matches the salient properties of equity returns and the risk-free rate, the dynamics of

*Itamar Drechsler is at Stern School of Business, New York University. I my grateful to my committee, Amir Yaron (Chair), Rob Stambaugh, and Stavros Panageas. I also thank Andy Abel, Ravi Bansal (discussant), Joao Gomes, Lars Hansen, Philipp Illeditsch, Jakub Jurek, Richard Kihlstrom, Feifei Li (discussant), Jun Liu, Nick Roussanov, Freda Song, Nick Souleles, Luke Taylor, Jessica Wachter, Paul Zurek, and seminar participants at Wharton, Chicago Booth School of Business, Columbia GSB, Princeton, NYU Stern, the University of Rochester, the 2009 Western Finance Association meeting (San Diego), the 2009 Stanford Institute for Theoretical Economics (SITE) workshop, and the 2010 American Finance Association meeting (Atlanta) for helpful comments. I thank Nim Drechsler and contacts at Citigroup and CSFB for options data. I am also indebted to the Editor, Cam Harvey, and to an anonymous referee for numerous comments that significantly improved the paper.

[1] I use the terms Knightian uncertainty, model uncertainty, and ambiguity interchangeably.

DOI: 10.1111/jofi.12068

conditional equity variance, and the moments of cash flows (consumption and dividends). The model is the first, to my knowledge, to successfully capture this broad set of features. It does this with a modest risk aversion of five, and as I explain below, a quantitatively reasonable level of model uncertainty.

Prior studies find that equity index option prices incorporate a considerable premium.[2] A good measure of this is the difference between option-implied expectations of return variance (which correspond to the VIX) and statistical (i.e., true) expectations of return variance (see Section II for details). The option-implied expectation is equivalent to the price of a variance swap, a contract that pays the buyer the realized variance over the term of the contract. Buying the swap provides the buyer with protection against high realized variance. Thus, the difference between the swap price and the true expectation of variance is the (variance) premium investors pay for this hedge. As I document below, at the one-month horizon the variance premium is large. Furthermore, its time series has predictive power for stock returns over horizons of a few months. A second prominent feature of index option prices is the shape of their implied volatility surface. Going back to at least Rubinstein (1994), academic studies have been concerned with understanding the sources of the volatility skew, which implies that otm put options are expensive relative to the benchmark Black–Scholes–Merton model (see, for example, Bates (2003)).

In this paper, I show that a desire by the representative investor to hedge his model uncertainty concerns naturally leads to a price premium in index options, and that the magnitude of the premium is sensitive to his level of model uncertainty. In the model, the representative investor has in mind a benchmark/reference model that represents his best estimate of the dynamics of economic fundamentals. However, due to model uncertainty, he is worried that this reference model is incorrect in some ways. Moreover, he wants his consumption and portfolio choice decisions to be robust to this possibility. He therefore considers the possibility that the true model actually lies in a set of alternative models.[3] To ensure that his model uncertainty concerns are reasonable, he considers alternative models that are statistically difficult to differentiate from the reference model and hence difficult to reject empirically. To make decisions that are robust across the set of alternative models, the investor optimally evaluates his decisions under whichever alternative model represents the "worst case." This worst-case alternative model is determined by perturbing (i.e., changing) the reference model in whichever ways are most damaging to the investor's utility, while remaining difficult to rule out based on historical data. The worst-case perturbations represent the potential modeling errors that most worry the investor.

[2] See, for example, Coval and Shumway (2001), Pan (2002), Bakshi and Kapadia (2003), Eraker (2004), and Carr and Wu (2009). Singleton (2006) contains a review of option pricing studies.

[3] My framework builds on the literature on robust control, pioneered by Hansen and Sargent. See, for example, Anderson, Hansen, and Sargent (2003), Hansen et al. (2006), and Hansen and Sargent (2008). The preferences are in the class of recursive multiple priors utility of Epstein and Schneider (2003).

Since the investor evaluates his decisions under the worst-case model, his concerns about model errors have an important impact on asset prices. For example, one of the investor's concerns is that his reference model overstates the mean of the persistent component in cash flow growth rates. Consequently, he decreases the price he is willing to pay for equity, which increases the expected equity return. In the calibrated model, the investor optimally allocates the biggest fraction of his uncertainty to concerns about the reference model's specification of jump shocks. Jump shocks are infrequent but relatively large shocks to the expected growth and growth volatility processes. The investor is concerned that the reference model underestimates their frequency or magnitude, which means they are amplified under the worst-case model. This increases the price the investor is willing to pay for variance swaps and helps generate a large variance premium and steep implied volatility skew. Stated another way, the investor knows that, if these model uncertainty concerns materialize, the resulting realized variance will be high. As index options will have a high payoff if this occurs, they provide a hedge to these model uncertainty concerns and the investor is therefore willing to pay a large premium for them. A further implication is that the size of the variance premium will vary to reflect fluctuations in the investor's level of uncertainty. This, in turn, will endow it with power to predict equity returns.

The framework constructed in this paper involves model uncertainty over a richer set of economic dynamics than have been possible in previous applications of robust control or ambiguity aversion. Specifically, the calibrated model incorporates uncertainty over a reference model whose dynamics include (i) a persistent component in expected cash flow growth, (ii) stochastic volatility, and (iii) jumps in the expected growth and growth volatility processes.[4] This rich set of dynamics is required for the model to be flexible enough to confront a large set of features of equity and options prices, and to allow model uncertainty to operate through multiple channels. The framework also extends previous models in allowing for time-variation in the representative investor's level of model uncertainty and by separating risk aversion and intertemporal elasticity of substitution (IES). Both of these features are important for the calibrated model to match the data. Despite this increased complexity, I provide analytical expressions for equity prices and the variance premium, which show the impact of uncertainty concerns on these quantities. Having analytical expressions also allows for precise calculation of option prices and the implied volatility surface.

The calibration of the model matches a broad set of features of cash flows and asset prices. These include "traditional" targets of equilibrium asset pricing models—the high equity premium, low risk-free rate, and excess volatility of equity returns relative to fundamentals. At the same time, the model

[4] Tractability is an issue when solving for equilibrium prices in models with uncertainty. Many financial applications focus on either log utility, which aids tractability, or i.i.d or single state variable environments. Some examples are Kleshchelski and Vincent (2007), Ulrich (2013), Brevik (2008), Trojani and Sbuelz (2008), Maenhout (2004), and Uppal and Wang (2003).

generates the aforementioned features of index option prices, which are not captured by leading equilibrium asset pricing models.[5] The calibrated model generates a variance premium that is large and positive, as in the data, and has predictive power for one- and three-month excess stock returns that is empirically consistent. Furthermore, the calibration matches the implied volatility surface, with an implied volatility skew that is very steep for one-month options and decays steadily as the horizon increases. Finally, the model also matches the dynamic properties of conditional equity variance, such as its persistence and volatility, and the higher moments of equity returns.

The calibration achieves its fit to the data using a level of model uncertainty under which the worst-case model is difficult to reject empirically; roughly speaking, it implies that one cannot reject at the 10% level the null hypothesis that the worst-case model generates the data. It is therefore reasonable for the investor to be concerned about this model. At the same time, the level of relative risk aversion used in the model is only five, which is much below the risk aversion levels used in most macro-finance asset pricing models.[6] If, however, uncertainty is turned off, then a much higher risk aversion of 13 is required to match the equity premium under the calibration, yet the variance premium generated in that case is too small.

The paper is organized as follows. Section I relates the paper to the literature. Section II discusses the data and empirical evidence. Section III presents the model. Section IV solves for the representative agent's equilibrium value function and worst-case model. Section V solves for equilibrium asset prices. Section VI calibrates the model and assesses its fit to the data. Section VII concludes. The Appendix contains technical material, while the Internet Appendix contains additional analysis and remaining technical details.[7]

I. Research Setting

This paper is part of a new literature that bridges the gap between research on equilibrium models of asset prices and studies of reduced-form option pricing models. The findings of the reduced-form studies are informative for equilibrium models, whose traditional set of targets—the equity premium and risk-free rate—are too limited to provide a full picture of the sources of risk and uncertainty impacting prices. In particular, a number of studies (e.g., Pan (2002), Eraker (2004), and Broadie, Chernov, and Johannes (2007)) find

[5] Du (2010) shows that the habit models of Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004) are inconsistent with the implied volatility skew of index options, while Backus, Chernov, and Martin (2011) and Du (2010) show that this is also the case for the rare disasters model of Barrow (2006). Below I document that the long-run risk models of Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2007) produce a variance premium that is much too small relative to the data.

[6] Barillas, Hansen, and Sargent (2009) highlight the result that, with model uncertainty, a desire for robustness allows for a reduction in the risk aversion coefficient.

[7] The Internet Appendix may be found in the online version of this article.

that jumps are important for describing the time series of equity returns and volatility, and that jump risk is significantly priced in options. The equilibrium model in this paper incorporates these features and is consistent with these findings. Recent studies also focus on the variance premium as an important object of interest, as it is in this paper. Broadie, Chernov, and Johannes (2009) argue that looking at the wedge between implied and realized volatility (essentially the variance premium) is the most informative way to analyze the premium in option prices. Todorov (2010) emphasizes the importance of jumps in explaining the dynamics and pricing of the variance premium, while Bollerslev, Gibson, and Zhou (2011) and Bollerslev, Tauchen, and Zhou (2009) find that their variance premium measures have significant predictive power for stock returns at horizons of a few months.

There are a few examples of equilibrium models that do target features of option prices. These include Benzoni, Collin-Dufresne, and Goldstein (2011), Du (2010), Eraker and Shaliastovich (2008), Liu, Pan, and Wang (2005; hereafter LPW), and Shaliastovich (2011). LPW is discussed below, while a discussion of these other papers and additional potential approaches is left to the Internet Appendix. None of these other models, however, considers the full set of equity, option, and variance features addressed by this paper, and in particular, none analyzes the variance premium.

LPW are the first to use model uncertainty to explain option prices in equilibrium. In their model, investors' uncertainty about rare events explains the skew in index option implied volatilities. In contrast to this paper, the model in LPW has i.i.d. dynamics and equates the volatility of consumption and returns. As a result, it cannot capture the "excess volatility" of returns, the dynamics of conditional variance, or return predictability by the variance premium. LPW's calibration is limited to the equity premium and volatility skew. They do not consider the variance premium or other moments of equity returns, the risk-free rate, or cash flows. In addition, jumps in LPW represent "rare disasters," that is, rare and very large jumps in the aggregate endowment, while in this paper jumps occur on average every one to two years and are moderate by comparison. In LPW, jumps enter into the *level* of cash flows, while the jumps here are in the expected *growth rate* of cash flows, so the level of cash flows evolves smoothly.[8] Finally, LPW do not quantify the amount of uncertainty implied by their calibration, while this quantity is an important part of this paper's calibration.

This paper is also related to Anderson, Ghysels, and Juergens (2009), who construct measures of the economy-wide level of (Knightian) uncertainty using survey forecasts and find that these measures predict stock returns in both the time series and the cross section. This paper also creates a proxy for economy-wide model uncertainty based on survey forecasts and shows that it closely tracks fluctuations in the variance premium.

[8] Benzoni, Collin-Dufresne, and Goldstein (2011) is the first paper to model jumps in the expected growth process.

Methodologically, this paper contributes to the literature applying model uncertainty to financial applications. Maenhout (2004) solves for the equity premium in an economy with a robust investor that has recursive utility. He exogenously modifies the canonical robust control model to induce wealth independence and analytical tractability. This exogenous modification is also used in Uppal and Wang (2003) and LPW. In contrast, wealth-independence arises endogenously in the framework of this paper.

This paper is also related to Drechsler and Yaron (2011), who build an extended long-run risks (Bansal and Yaron (2004)) model with jump shocks that captures the size and predictive power of the variance premium. They do not consider model uncertainty, a central driver of asset prices in this paper, and do not examine the option-implied volatility surface or the connection between option prices and measures of model uncertainty. While in this paper time-variation in the perceived risk of jumps also drives fluctuations in the variance premium, it now arises from the combination of reference-model jump risks and fluctuations in model uncertainty fears. Moreover, since model uncertainty endogenously amplifies concerns about jump shocks, less is needed in terms of physical jumps, and the model is able to capture the equity premium and option prices with a significantly lower risk aversion of five. Finally, the model in this paper generates stochastic return volatility from two channels, stochastic cash flow volatility and stochastic uncertainty. This makes the variance premium imperfectly correlated with conditional volatility and a better predictor of equity returns than overall variance.

Finally, a number of papers prominently feature uncertainty in an asset pricing context, though under the Bayesian approach. Pástor (2000), Pástor and Stambaugh (2000), and Avramov (2004) use a Bayesian approach to asset allocation. Avramov (2002) and Pástor and Stambaugh (2011) assess return predictability, Brennan and Xia (2001) model how learning about predictability impacts asset allocation, and Veronesi (2000) shows that Bayesian filtering on the part of investors can induce interesting dynamics in asset prices. These papers analyze different problems from the one studied here, and none of them constructs an equilibrium model of prices based on preferences and fundamentals, assuming instead an exogenous process for returns and/or the pricing kernel.

II. Data

The variance (risk) premium is defined as the difference between the risk neutral and physical (i.e., true/statistical) expectations of total stock market return variance for a given horizon. I focus on a one-month variance premium. Hence, the one-month variance premium at time t, denoted vp_t, equals $E_t^Q[\int_t^{t+1}(d\ln R_{m,s})^2] - E_t[\int_t^{t+1}(d\ln R_{m,s})^2]$, where $\ln R_{m,s}$ is the (log) return on the market and Q denotes the risk-neutral probability measure.

Demeterfi et al. (1999) and Britten-Jones and Neuberger (2000) show that, when the underlying asset price is continuous, the risk-neutral expectation of

variance for an asset is given by the value of a portfolio of European calls on that asset. This follows by demonstrating a trading strategy that uses a portfolio of options to replicate a variance swap on the asset. In other words, the strategy's payoff is the asset's realized variance. The risk-neutral expectation of this payoff then follows from the cost of the strategy, which is the cost of the option portfolio. Jiang and Tian (2005) and Carr and Wu (2009) extend this result to the case in which the asset follows a jump-diffusion. Note that this approach is "model-free" since it does not depend on any particular option pricing model. The Chicago Board Options Exchange (CBOE) uses this model-free method to calculate the VIX index as the risk-neutral expectation of variance for the S&P 500 over the subsequent 30 days. I obtain closing values of the VIX from the CBOE and use it as my measure of risk-neutral expected variance. More precisely, since the VIX index is reported in annualized "vol" terms, I square it to put it in "variance" space and divide by 12 to get a monthly quantity. I refer to the resulting series as squared VIX. For each month, I use the value of the VIX on the last day.

To measure vp_t one also needs a conditional forecast of total return variation under the physical measure. Measures of the realized variance of the market for a given month can be obtained by summing up the squared five-minute log returns on the S&P 500 futures and on the S&P 500 index over the whole month.[9] I obtain the high frequency data used to construct these realized variance measures from TICKDATA. To get a conditional forecast series, I use the one-step-ahead forecasts from a simple time-series model. I project the futures realized variance on the value of the squared VIX at the end of the previous month and on a lagged realized variance measure. The one-step-ahead forecast series from this regression serves as the proxy for the conditional expectation of total return variance under the physical measure. The difference between the risk neutral expectation, given by the squared VIX, and the contemporaneous conditional forecast gives the time series of one-month variance premium estimates. The intuition that the variance premium is a premium paid for insurance (against high realized variance) implies that it should be nonnegative month-by-month. That is, the physical expectation of variance should be less than the risk-neutral one. For 239 out of 240 months, this is not a problem since in 238 months the variance premium estimate is positive, while one month has a tiny negative value and can be considered zero. I impose the theoretical restriction on the remaining month by truncating the variance forecast proxy at the value of the squared VIX, implicitly truncating the variance premium estimate to zero. The impact of this truncation on the results, which is small,

[9] Estimating realized variance for highly liquid assets via the summation of high-frequency squared returns has become the standard method in practical applications (see, for example, Andersen et al. (2001), Andersen, Bollerslev, and Diebold (2007), and Bollerslev, Gibson, and Zhou (2011)). A five-minute sampling frequency provides a good balance between the improved estimation precision afforded by a finer sampling of the data and the greater measurement noise introduced into such increasingly finer sampling due to bid-ask spreads, price discreteness, and other microstructure issues (see, for example, Hansen and Lunde (2006)).

Table I
Summary Statistics

This table presents summary statistics for the squared VIX, realized variance, and variance premium. The sample is monthly. The columns labeled 2007.6 present estimates for the pre-crisis sample, 1990.1 to 2007.6, while the columns labeled 2009.12 present estimates for the full sample, 1990.1 to 2009.12. VIX^2 is the value of the CBOE's VIX index squared and divided by 12 to convert it into a monthly quantity. The value of VIX^2 for a particular month is the last observation of that month. Realized variance corresponds to the series of monthly realized variances, measured as the sum over a month of the squared five-minute log returns on the S&P 500 futures. Variance Premium corresponds to the measure of the one-month variance premium, constructed as the difference between VIX^2 and a forecast of next month's realized variance, as described in the text. The time series of this measure is used to calculate all of the variance premium's summary statistics except for the mean, which is estimated nonparametrically as the sample average of VIX_t^2 − realized variance_{t+1}.

	VIX^2		Realized Variance		Variance Premium	
	2007.6	2009.12	2007.6	2009.12	2007.6	2009.12
Mean	33.08	39.32	22.02	28.82	11.05	10.55
Median	24.55	30.88	14.10	16.71	7.55	7.96
Std.-Dev.	23.97	36.58	22.43	47.31	7.63	8.47
Maximum	163.4	298.9	142.5	566.9	58.84	58.84
Skewness	2.02	3.38	2.67	7.24	2.43	2.19
Kurtosis	8.98	19.21	11.49	74.75	12.34	9.67
AC(1)	0.79	0.82	0.64	0.65	0.61	0.61

is indicated below. The Internet Appendix provides additional details on the variance forecast proxy.

The data series for the VIX and realized variance measures covers the period January 1990 to December 2009. The main limitation on the length of the sample is the VIX, whose time series begins in January 1990. Daily data on the implied volatility surface of S&P 500 options, which is used in later analysis, is obtained from Citigroup and covers October 1999 to June 2008. Data on per-capita consumption of nondurables and services is taken from NIPA. The per-share dividend series for the stock market is constructed from CRSP by aggregating dividends paid by common shares on the NYSE, AMEX, and NASDAQ. Dividends are adjusted to account for repurchases as in Bansal, Dittmar, and Lundblad (2005).

Table I provides summary statistics for the VIX, futures realized variance, and the variance premium measure (VP). Throughout the paper, I report results for both the full sample (January 1990 to December 2009) and a sample that ends just before the first wave of the financial crisis began (January 1990 to June 2007). It is well known that there were very large spikes in volatility during the crisis, especially in October and November of 2008. This has a big effect on the estimates of some second- and higher-order moments, particularly those related to variance. The pre-crisis estimates provide an interesting

comparison across samples and allow one to assess the data when the influence of the outlier months is excluded.[10]

The table shows that the mean of the variance premium is sizable in comparison with those of the squared VIX and realized variance. Between a quarter and a third of the risk-neutral expectation of variance is a premium. More precisely, it represents an insurance premium, since the price of the variance swap is greater than its average payoff. As the table shows, the variance premium is also quite volatile. Note further that all three variance-related series display significant deviations from normality. The mean to median ratio is large, the skewness is positive and greater than zero, and the kurtosis is clearly much larger than three. Across the two samples the variance premium statistics are similar, though in the full sample the mean is a bit lower, the standard deviation a bit higher, and the kurtosis a bit lower but still very high.[11] There are bigger changes across the samples in the statistics of the other two series. In particular, the standard deviations and higher moments all increase dramatically, due to the tremendous spike in volatility in the fall of 2008 and the first quarter of 2009.[12]

Table II shows return predictability regressions. There are two sets of columns with regression estimates. The first set shows OLS estimates and the second provides estimates from robust regressions, which downweight the influence of outliers on the estimates, providing a check that they do not drive the results. The robust regression R^2s reported are calculated as the ratio of the variance of the regression forecast to the variance of the dependent variable, which corresponds to the usual R^2 calculation in the case of OLS. The first two regressions are one-month-ahead forecasts using the variance premium as a univariate regressor, while the third forecasts one quarter ahead. The quarterly return series is overlapping. The last two specifications add the price–earnings ratio, a commonly used variable for predicting returns. As a univariate regressor, the variance premium can account for about 1.5% to 4.9% of the monthly excess return variation. The multivariate regressions lead to a further increase in the R^2. In conjunction with the price–earnings ratio, the in-sample R^2 can increase to over 10%. Note that in all cases the variance premium enters with a significant positive coefficient, which will be shown to

[10] The starting point of the crisis is often marked by the June 2007 collapse of two Bear Stearns hedge funds and the August 2007 "quant crisis." However, the results are similar if the pre-crisis sample is extended until August 2008.

[11] The truncation of the negative VP month only has a noticeable impact on the kurtosis and skewness of the full-sample VP. Without truncation, the kurtosis of VP rises to 10 and the skewness is 1.60.

[12] I also estimate moments for VP for the sample 1986 to 2009 and find that they are broadly similar to those for the post-1990 sample, with an estimated mean and standard deviation of the variance premium of 9.44 and 12.77, respectively. I obtain a proxy for VIX for the pre-1990 sample by using VXO (the "old" VIX), which is available from 1986. I regress VIX on VXO and a constant within the 1990 to 2009 sample and use the estimated relationship to obtain fitted values for VIX for the pre-1990 sample. For October 1987, I also replace the high-frequency futures realized variance with daily realized variance because of the well-documented breakdown in futures trading that occurred during the stock market crash.

Table II
Return Predictability by the Variance Premium

This table presents return predictability regressions. The sample is monthly. The top panel presents the estimates for the pre-crisis sample, 1990.1 to 2007.6, while the bottom panel shows the estimates for the full sample, 1990.1 to 2009.12. Reported t-statistics for OLS are Newey–West (HAC) corrected. VP is the variance premium. P/E is the price–earnings ratio for the S&P 500. The dependent variable is the log excess return (annualized and in percent) on the S&P 500 index over the following one and three months, as indicated. The three-month returns series is overlapping. Robust Reg. denotes estimates from robust regressions using a bisquare weighting function. The reported robust regression pseudo-R^2s are calculated as the ratio of the variance of the regression forecast to the variance of the dependent variable, which corresponds to the usual R^2 calculation in the case of OLS.

	Regressors		OLS			Robust Reg.		
	X1	X2	β_1	β_2	$R^2(\%)$	β_1	β_2	$R^2(\%)$
Pre-Crisis Sample (1990.1 to 2007.6)								
r_{t+1}	VP_t		0.74		1.44	1.08		3.05
	(t-stat)		(2.18)			(2.71)		
r_{t+1}	VP_{t-1}		1.32		4.48	1.22		3.88
	(t-stat)		(3.96)			(3.04)		
r_{t+3}	VP_t		0.87		6.15	0.88		6.29
	(t-stat)		(3.53)			(4.26)		
r_{t+1}	VP_t	$\log(\mathrm{P/E})_t$	1.31	−47.43	8.08	1.69	−48.85	10.22
	(t-stat)		(2.89)	(−3.07)		(4.14)	(−4.28)	
r_{t+1}	VP_{t-1}	$\log(\mathrm{P/E})_t$	2.06	−56.67	13.64	1.95	−56.34	12.82
	(t-stat)		(4.73)	(−3.51)		(4.71)	(−4.86)	
Full Sample (1990.1 to 2009.12)								
r_{t+1}	VP_t		0.80		1.72	1.40		5.24
	(t-stat)		(2.25)			(3.99)		
r_{t+1}	VP_{t-1}		1.38		5.10	1.46		5.68
	(t-stat)		(3.97)			(4.11)		
r_{t+3}	VP_t		0.99		6.93	1.11		8.62
	(t-stat)		(3.98)			(5.72)		
r_{t+1}	VP_t	$\log(\mathrm{P/E})_t$	1.37	−25.03	4.69	1.86	−23.74	7.30
	(t-stat)		(3.70)	(−1.49)		(4.67)	(−2.86)	
r_{t+1}	VP_{t-1}	$\log(\mathrm{P/E})_t$	2.17	−33.65	10.34	2.06	−30.79	9.20
	(t-stat)		(5.32)	(−2.20)		(5.09)	(−3.63)	

be consistent with the theory in this paper. In addition, the robust regression estimates agree in both magnitude and sign with the OLS estimates. Comparing across the two samples, the coefficient estimates and predictive R^2s do not change greatly.[13]

[13] If the VP series is *not* truncated at zero (which affects only the full sample) then the predictive power is actually a bit greater. I also split the full sample into equal-length subperiods, January 1990 to December 1999, and January 2000 to December 2009, and estimate the predictability regressions within the subsamples. The OLS estimates for the coefficient on the variance premium in

Table III

Return Predictability by the Variance Premium: Rolling Regressions

This table presents return predictability regressions based on a VP series constructed from variance forecasts taken from rolling regressions. The sample is monthly. The columns labeled 2007.6 present estimates for the pre-crisis sample 1990.1 to 2007, while the columns labeled 2009.12 are for the full sample, 1990.1 to 2009.12. The first 24 data points (months) are used to initialize the rolling regression for the variance forecast, so the effective sample for return prediction begins in 1992.1. Reported t-statistics are Newey–West (HAC) corrected. P/E is the price–earnings ratio for the S&P 500. The dependent variable is the log excess return (annualized and in percent) on the S&P 500 index over the following one and three months, as indicated. The three-month returns series is overlapping.

	Regressors		OLS (2007.6)			OLS (2009.12)		
	X1	X2	β_1	β_2	$R^2(\%)$	β_1	β_2	$R^2(\%)$
r_{t+1}	VP_t		0.71		2.21	0.73		1.86
	(t-stat)		(2.94)			(2.39)		
r_{t+1}	VP_{t-1}		1.00		4.35	1.28		5.69
	(t-stat)		(3.85)			(4.16)		
r_{t+3}	VP_t		0.70		6.79	0.92		7.63
	(t-stat)		(3.38)			(3.98)		
r_{t+1}	VP_t	$\log(\mathrm{P/E})_t$	1.43	−60.91	12.61	1.08	−21.21	4.29
	(t-stat)		(5.84)	(−4.77)		(3.98)	(−1.27)	
r_{t+1}	VP_{t-1}	$\log(\mathrm{P/E})_t$	1.91	−70.94	17.86	1.79	−29.00	10.10
	(t-stat)		(4.79)	(−4.89)		(4.45)	(−1.69)	

As an additional robustness check, I construct the VP series using one-step-ahead variance forecasts from regressions estimated on a rolling basis using only past data, rather than the whole sample. The first 24 months are used to initialize the rolling regression estimates, so this VP series begins in January 1992. Table III repeats the return predictability regressions using this VP series. A comparison with Table II shows that the results are largely unchanged.

A. An Empirical Proxy for the Level of Model Uncertainty

Figure 1 graphs an empirical proxy for the economy-wide level of model uncertainty. I measure the level of model uncertainty by the dispersion in the set of forecasts of next quarter's real GDP growth, from the Philadelphia Fed's Survey of Professional Forecasters (SPF). The dispersion is calculated simply as the standard deviation in the growth forecasts, which are reported near the beginning of every quarter. I take the dispersion as a proxy for the size of the alternative set of reasonable models that the representative investor considers at each point in time. This measure maps nicely to the way time-varying uncertainty is formally modeled below. Moreover, information asymmetry should not be an important issue for forecasting an aggregate quantity such as GDP, since the relevant information is widely reported and discussed in the public

the one-month-ahead regressions are 1.07 for the earlier subperiod and 0.75 for the later subperiod. The difference of 0.32 is *not* statistically significant, with a t-statistic of 0.49.

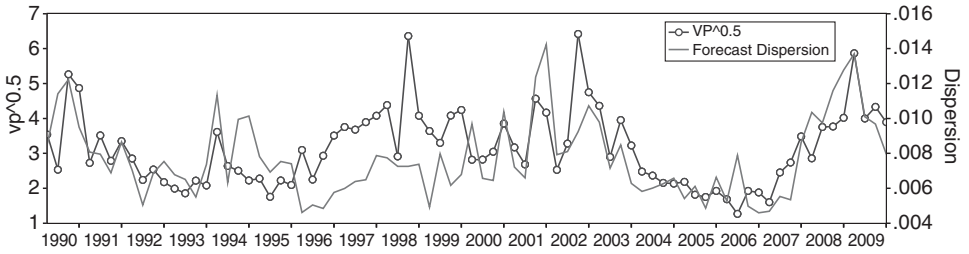

Figure 1. Forecast dispersion and variance premium. The figure plots the standard deviation in forecasts of next quarter's real GDP growth from the Survey of Professional Forecasters (SPF) versus the square root of VP at the end of the previous quarter. The sample is 1990Q1 to 2009Q4. The correlation of the two series is 0.54 with a robust standard error of 0.14.

domain. It is therefore natural to interpret different forecasts as the product of different economic models. The forecast dispersion captures the size of the cross section of model forecasts that the representative investor considers (i.e., the alternative set).

The figure plots the quarterly dispersion measure along with the square root of VP at the end of the previous quarter. Taking the square root puts both measures in standard deviation terms. The comovement between the two series is striking. Their correlation is 0.54, with a standard error of 0.14. The two series tend to spike around the same time, particularly in 1990 to 91, 1994, 2001 to 2002, and the recent crisis. Note that the dispersion measure begins to rise from a low level in the third quarter of 2007, when the first signs of the economic crisis emerge. One exception to their strong comovement is the financial crisis of 1998, which caused a sharp spike in the options market without a corresponding strong increase in forecast dispersion. Perhaps this is due to the short-lived nature of this episode, which appeared to be largely international in nature and not directly related to domestic economic events. Overall, the figure provides support for a link between time-variation in model uncertainty and the premium in index option prices, a relationship predicted by the model developed below.[14]

III. Model Framework

The setting for the model is an infinite-horizon, continuous-time exchange economy with a representative investor who has utility over consumption streams. This investor has in mind a benchmark or *reference* model of

[14] Ben-David, Graham, and Harvey (2010) ask CFOs for the 10th and 90th percentiles of their predictive distribution for next year's S&P 500 return. The average 10th percentiles series, which is informative about CFOs' concerns of reasonable worst-case return outcomes, has a correlation of −0.63 and −0.76 with the dispersion series and its one-period lead, respectively, over the sample period 2000Q2 to 2010Q1. The relationship between these two series, which are drawn from different sources, is evidence that the uncertainty fluctuations measured by the dispersion series are not idiosyncratic to professional forecasters, but reflect economy-wide variation.

the economy that represents his best estimate of the economy's dynamics. However, the investor does not fully trust that his model is correct, causing him to worry that the true model lies in a set of *alternative* models that are difficult for him to reject empirically. These alternative models are "close" to the reference model because they are difficult to distinguish from it statistically. The investor guards against model uncertainty by making consumption and portfolio choice decisions that are robust across the set of alternative models. This is equivalent to evaluating future prospects under whichever is the *worst-case* model in the alternative set. I now formalize this setup in detail.

A. *The Reference Model*

Let $Y_t \in \mathbb{R}^n$ denote the vector of state variables. It follows a continuous-time affine jump-diffusion:[15]

$$dY_t = \mu(Y_t)dt + \Sigma(Y_t)dZ_t + \xi_t \cdot dN_t, \tag{1}$$

where Z_t is a Brownian motion in $\mathbb{R}^n$, $\Sigma(Y_t)$ is a matrix, and $\mu(Y_t)$, ξ_t, and N_t are vectors. The term $\xi_t \cdot dN_t$ denotes component-wise multiplication of the jump sizes in the random vector ξ_t and the increments in the vector of Poisson (counting) processes N_t. The Poisson arrivals are conditionally independent and arrive with a time-varying intensity given by the vector $l_t \in \mathbb{R}^n$. The jump sizes in ξ_t are assumed to be i.i.d. Let $\psi(u)$ denote the vector that stacks the moment-generating functions of the jump sizes; its kth component is $\psi_k(u_k) = E[\exp(u_k\xi_k)]$, which completely characterizes the distribution of the jump size ξ_k. For convenience, let log consumption and dividends, $\ln C_t$ and $\ln D_t$, be in Y_t, and let δ_c and δ_d be vectors that select them from Y_t, that is, $\ln C_t = \delta_c' Y_t$ and $\ln D_t = \delta_d' Y_t$. Although $\ln C_t$ is in Y_t, I maintain the standard assumption that the dynamics of Y_t do not depend on $\ln C_t$, so that equilibrium quantities are homogeneous in the level of consumption.

The drift, diffusion, and jump intensity functions have an affine structure. The drift is given by

$$\mu(Y_t) = \mu + \mathcal{K}Y_t,$$

where $\mu \in \mathbb{R}^n$ and $\mathcal{K} \in \mathbb{R}^{n\times n}$. The diffusion covariance matrix $\Sigma(Y_t)\Sigma(Y_t)'$ is assumed to have a block-diagonal structure, where each block is affine. This structure is necessitated by the model uncertainty setup, as explained below. Towards that end, let $Y_t = (Y_{1,t}', Y_{2,t}')'$ be a partition of the state vector. This partition determines which subset of the dY_t dynamics the investor is uncertain about. The diffusion covariance matrix takes the block-diagonal form

$$\Sigma(Y_t)\Sigma(Y_t)' = \begin{bmatrix} \Sigma_{1,t}\Sigma_{1,t}' & 0 \\ 0 & \Sigma_{2,t}\Sigma_{2,t}' \end{bmatrix},$$

[15] The notation used here for the jump-diffusion specification is similar to that in Duffie, Pan, and Singleton (2000) and Eraker and Shaliastovich (2008).

where the upper block corresponds to $Y_{1,t}$ and the lower block to $Y_{2,t}$. The quantity $\Sigma_{1,t}\Sigma_{1,t}'$ can take the general affine form: $\Sigma_{1,t}\Sigma_{1,t}' = h + \Sigma_i H_i Y_{t,i}$. Let q_t denote a particular state variable in Y_t. This variable will appear repeatedly throughout the model and has the important role of governing variation in the investor's level of uncertainty. I assume that

$$\Sigma_{2,t}\Sigma_{2,t}' = H_q q_t^2.$$

Finally, I let the jump intensity vector take the form $l_t = l_1 q_t^2$, where l_1 is a vector in $\mathbb{R}^n$.[16]

The partition of Y_t describes which subset of the dynamics dY_t the investor is uncertain about. I make the investor uncertain about only the dynamics of $dY_{2,t}$, which contains the dynamics of the variables the investor feels are difficult to detect or estimate.[17] Thus, the specification above makes q_t drive the volatility of shocks about which there is uncertainty. This specification is motivated by considerations of both tractability and empirical plausibility, as discussed below.

B. Alternative Models

I now describe the set of alternative models contemplated by the investor. The idea is to first consider the most general set of alternative dynamics possible, and then restrict the alternatives to a subset that are statistically close to the reference model. A model is defined by its probability measure. Let P be the probability measure associated with the reference model (1). An alternative model of dY_t is defined by a probability measure $P(\eta)$, where η_t is a process for the likelihood ratio (Radon-Nikodym derivative) of $P(\eta)$ with respect to P. The investor considers alternative models of the form

$$dY_t = [\mu(Y_t) + \Sigma(Y_t)h_t]dt + \Sigma(Y_t)dZ_t^\eta + \xi_t^\eta \cdot dN_t^\eta. \tag{2}$$

In addition, denote the moment-generating function under $P(\eta)$ by $\psi^\eta(u)$. Equation (2) looks somewhat like (1), but there are a number of changes. These changes are referred to as perturbations to the reference model, and the resulting model is called the perturbed model. The process Z_t^η is a Brownian motion, but now under the probability measure $P(\eta)$. The drift of dY_t is now perturbed by the new term $\Sigma(Y_t)h_t$, which depends on the vector $h_t \in \mathbb{R}^n$. Since there is model uncertainty only over the dynamics of $dY_{2,t}$, I impose $h_t = [0, h_{2,t}']'$, where the zero and $h_{2,t}$ vectors have the dimensions of $Y_{1,t}$ and $Y_{2,t}$, respectively. The block-diagonal structure of the diffusion covariance matrix then implies that only the drift of $Y_{2,t}$ is perturbed. The vector process $h_{2,t}$ is left completely free, which means the perturbations to the drift of $dY_{2,t}$ are unrestricted. If $h_t = 0$, we are back to the reference model.

[16] It is also possible to partition the jump intensity vector and let the intensity for one partition have the general affine form. Since this generality is not needed in what follows, I omit it to reduce notational complexity.

[17] We can, of course, have $Y_{2,t} = Y_t$.

The other set of perturbations change the jump intensity and jump size distribution under $P(\eta)$. This is indicated by the explicit dependence on η_t in the jump term $\xi_t^\eta \cdot dN_t^\eta$ and in $\psi^\eta(u)$. Unlike the drift perturbation, the potential jump perturbations are not completely unrestricted, but they do encompass a wide range of alternatives. Consider first the jump intensity. Under $P(\eta)$ it is given by

$$l_t^\eta = \exp(a) l_t,$$

where a is a scalar parameter that amplifies or diminishes the jump intensity. For the jump sizes, I consider two specific jump size distributions, which are the ones used in the model calibration: (i) normally distributed jumps: $\xi_j \sim \mathcal{N}(\mu, \sigma^2)$, and (ii) gamma distributed jumps: $\xi_j \sim \Gamma(k, \theta)$, where k and θ are the gamma shape and scale parameters, respectively. Under $P(\eta)$, the jump size distributions change so that

$$\xi_j^\eta \sim \mathcal{N}(\mu + \Delta\mu, \sigma^2 s_\sigma), \qquad \xi_j^\eta \sim \Gamma\left(k, \frac{\theta}{1-\theta b}\right).$$

For the normal distribution, the mean is shifted by an amount $\Delta\mu$, while the variance is scaled by s_σ. For the gamma distribution, the scale parameter is increased or decreased depending on the sign of b. Changing the scale parameter changes both the mean and the variance of the gamma distribution. Note that, when $\Delta\mu = b = a = 0$ and $s_\sigma = 1$, we are back to the jump distributions of the reference model.

Each alternative dynamics given by an instance of (2) corresponds to a specific process η_t. Knowing this likelihood ratio process is useful since, as shown in the next section, it gives an elegant way to restrict the set of alternative models to the set that are statistically difficult to distinguish from the reference model. By Girsanov's theorem for Itô-Lévy processes, $\eta_t = \eta_t^{dZ}\eta_t^J$, where η_t^{dZ} changes the probabilities for Z_t and η_t^J changes the probabilities for the jumps. The process for η_t^{dZ} is defined by the SDE $\frac{d\eta_t^{dZ}}{\eta_t^{dZ}} = h_t^T dZ_t$ and $\eta_0^{dZ} = 1$. By Girsanov's theorem, $Z_t^\eta = Z_t - \int h_t dt$ is then a Brownian motion under $P(\eta)$, which accounts for the source of the drift perturbation. The construction of η^J and further details are in the Internet Appendix.

In summary, the alternative dynamics the investor contemplates are determined by the sets of h_t, a, $\Delta\mu$, s_σ, and b that he considers at time t.[18] Moreover, the process η_t can be used to restrict consideration to the subset of these dynamics that are difficult to distinguish from the reference dynamics.

[18] In other words, these sets determine the investor's multiple priors over one-step-ahead probabilities. The investor's behavior falls within the multiple priors framework axiomatized by Epstein and Schneider (2003). As Epstein and Schneider (2003) show, when beliefs are built up as the product of one-step-ahead probabilities, the investor's decision-making is guaranteed to be dynamically consistent.

C. The Size of the Alternative Set

A commonly used measure of the statistical "distance" between a model and a reference model is its *relative entropy*. Relative entropy is directly related to statistical detection and is defined in terms of the process η_t. I restrict the investor's set of alternative models to those that are statistically close by putting an upper bound on the growth rate of alternative models' relative entropy.[19]

The growth in entropy of $P(\eta)$ relative to P between time t and $t + \Delta t$ is defined as $H(t, t + \Delta t) = E_t^{\eta}[\ln \eta(t + \Delta t)] - \ln \eta(t)$. Thus, $R(\eta_t) \stackrel{\text{def}}{=} \lim_{\Delta t \to 0} \frac{H(t,t+\Delta t)}{\Delta t}$ gives the instantaneous *growth rate* of relative entropy at time t. It is instructive to look at this quantity for the diffusion perturbation. A standard calculation (see Appendix A, shows that $R(\eta_t^{dZ}) = \frac{1}{2} h_t' h_t$. This simple expression says that the relative entropy growth rate at time t is just half the norm of the h_t vector. Hence, for $h_t = 0$ (the reference model), the rate is zero. As h_t increases so does $R(\eta_t^{dZ})$, highlighting the tight link between relative entropy and the distance between $P(\eta)$ and P. Moreover, this indicates that the set of alternative models can be restricted by placing an upper bound on the permissible relative entropy growth rates of these models.

Since $\eta_t = \eta_t^{dZ} \eta_t^{J}$, $R(\eta_t)$ is just the sum of $R(\eta_t^{dZ})$ and $R(\eta_t^{J})$. Moreover, $R(\eta_t^{J})$ is just the sum of the relative entropies for the individual jump perturbations. Appendix A derives the relative entropies for the normal and gamma jump perturbations and gives an expression for the total relative entropy growth rate, $R(\eta_t)$. As Appendix A shows, $R(\eta_t)$ is expressed in terms of $(h_t,\ a,\ \Delta u,\ s_\sigma,\ b)$.

To exploit the link between entropy and statistical proximity I define the set of alternative models considered by the investor to be all the alternative models for which $R(\eta_t)$ is less than a given upper bound. Intuitively, if the relative entropy growth rate of a model is below the bound, then distinguishing it from the reference model is reasonably difficult and there is a reasonable concern that this alternative model (and not the reference model) is the true data generating process. The bound on entropy therefore determines the level of model uncertainty. A large bound represents high uncertainty, since alternatives that are statistically further from the reference model will fall below the bound. In that case, the investor has little confidence in the reference model. At the other extreme, a bound of zero on $R(\eta_t)$ means the alternative set is empty and the investor has full confidence in the reference model.

To model time-varying uncertainty, the bound on $R(\eta_t)$ is allowed to vary over time based on the value of q_t^2, the variable controlling variation in the level of uncertainty. Hence, the alternative set of models considered at time t is defined by

$$\left\{\eta_t : R(\eta_t) \leq \varphi q_t^2\right\}, \tag{3}$$

[19] The use of entropy and the link to statistical detection is due to Hansen and Sargent (see Hansen and Sargent (2008)). The approach used here for time-varying uncertainty is used in Trojani and Sbuelz (2008) in a pure diffusion setting.

where $\varphi > 0$ is a constant. Since $q_t^2 > 0$, the bound is always positive. Without loss of generality, I normalize the process for q_t^2 so that $E[q_t^2] = 1$. Then the unconditional mean of the bound is simply equal to φ, while variation in the bound is due to q_t^2. The constant φ is part of the investor's preferences. If $\varphi = 0$ then the investor has full confidence in the reference model, while increasing the value of φ expands the alternative set to include models that are statistically further away from the reference model. While φ determines the investor's average level of uncertainty, q_t^2 controls variation in uncertainty over time. When q_t^2 increases, the investor is more uncertain and worries also about dynamics that are further away from that of the reference model.

The investor is worried about alternative models because he is concerned that one of them generates the data while he has mistakenly concluded that it comes from the reference model. This is called making a model detection error. If the statistical distance between an alternative model and the reference model is large, it is easy to distinguish between them using the data and the probability of detection error is small. As the distance from the reference model decreases, the probability of detection error increases. In calibrating the model, the value of φ is set to give a specific probability of detection error for the worst-case alternative model. The idea is that if the probability of making a model detection error is reasonably high for the worst-case model, then it is sufficiently difficult to distinguish it from the reference model to warrant concern that it may actually be the true model. Further details are give in Section V and in the Internet Appendix.

D. Utility Specification

In addition to incorporating model uncertainty, I want to allow for separation of the representative investor's IES and risk aversion, since these represent distinct properties of preferences. To that end, I assume that, *conditional* on a particular probability belief (i.e., model), the investor's preferences over consumption streams are determined by the recursive preferences of Epstein and Zin (1989), specified in continuous time as in Duffie and Epstein (1992). Hence, let J_t denote the representative investor's value function and let $f(c_s, J_s)$ be the normalized aggregator of consumption and continuation value in each period. The investor's utility is then given by

$$J = \max_{\{c_s\}} \min_{P(\eta)} E_0^{\eta}\left[\int_0^{\infty} f(c_s, J_s)ds\right],$$

where E^{η} denotes expectation taken under the probability measure $P(\eta)$. The minimization is taken over the set of probabilities given by the reference and alternative models. The investor expresses his model uncertainty and desire for robustness by evaluating his future prospects under the worst-case (i.e., minimizing) model within his set of alternatives, given his consumption choices. As usual, market clearing implies that the representative investor consumes the aggregate consumption, so that $c_s^* = C_s$. Epstein and Schneider (2003)

show that rectangularity of beliefs implies that J_t solves the Hamilton–Jacobi–Bellman (HJB) equation:

$$0 = \min_{P(\eta_t)} f(C_s, J_s) + E_t^{\eta}[dJ] \tag{4}$$

$$\text{s.t.} \quad R(\eta_t) \leq \varphi q_t^2.$$

The solution to this equation gives both J_t and the worst-case model, η_t^*.

Under the Duffie–Epstein–Zin parameterization of recursive preferences, f is given by

$$f(C, J) = \delta \frac{\gamma}{\rho} J \left[\frac{C^{\rho}}{\gamma^{\frac{\rho}{\gamma}} J^{\frac{\rho}{\gamma}}} - 1 \right], \tag{5}$$

where δ is the rate of time preference, γ is $1 - \text{RRA}$ (i.e., one minus the investor's relative risk aversion), and $\rho = 1 - \frac{1}{\psi}$, where ψ is the IES. In the special case where $\gamma = \rho$ (i.e., relative risk aversion equals $1/\psi$), the aggregator reduces to the additive power utility function.

IV. Solution

To solve for J_t and the worst-case model, one can proceed as follows. Expand the expectation in (4) using Ito's lemma and evaluate the η-expectation using the alternative dynamics from equation (2). Rewrite (4) as a Lagrangian with Lagrange multiplier λ_t on the time-t (relative entropy growth) constraint. Take first-order conditions with respect to the perturbation parameters $(h_t, a, \Delta u, s_\sigma, b, \ldots)$ and λ_t. Then conjecture a guess for J_t and verify that it solves the system of first-order conditions and the HJB equation. The solution for J is now discussed. Further details, including the first-order conditions, are left to Appendix B and Appendix C.

A. Equilibrium Value Function

From (5) it follows that J_t is homogeneous of degree γ in the level of consumption C_t. For notational convenience in what follows, let $Y_t = [\ln C_t, \tilde{Y}_t']'$. The value function can then be written as

$$J(Y_t) = \exp(\gamma g(\tilde{Y}_t)) \frac{C_t^{\gamma}}{\gamma} = \frac{\exp(\gamma g(\tilde{Y}_t) + \gamma \ln C_t)}{\gamma}, \tag{6}$$

where $g(\tilde{Y}_t)$, a function of $\tilde{Y}_t$, must be determined. Appendix B substitutes this expression for J_t into (4) and derives the resulting expression. In general, that equation does not have an exact analytical solution. However, I find an analytical approximation to the solution using a log-linear approximation to the consumption-wealth ratio around its (endogenous) unconditional mean. This approximation has been used successfully in the portfolio choice literature (see

Campbell et al. (2004)). The approximation is exact for $\psi = 1$ (and any value of γ) and remains accurate for a wide range of values around one.[20] It is provided by the following proposition.

PROPOSITION 1: *The solution to (5) for $\psi = 1$, or for $\psi \neq 1$ when the log-linear approximation is applied, is given by (6) and*

$$g(\tilde{Y}_t) = A_0 + A'\tilde{Y}_t, \tag{7}$$

where the scalar A_0, the vector A, and the worst-case model perturbations are determined by the solution to a system of equations given in Appendix C.

Note that the vector $\hat{A} = [1, A']'$ gives the sensitivities of the value function to the state variables Y_t (the 1 corresponds to C_t). Hence, the sign and magnitude of $\hat{A}(i)$ (the ith component) determines whether the value function increases or decreases in the value of the state variable $Y(i)$, and how strongly. As Proposition 1 indicates, the solution to the system of equations that gives the value function also gives the solution for the worst-case model.

B. Worst-Case Model

The worst-case model perturbations satisfy the first-order conditions for the minimization in (4). Their general form is illustrated by the first-order condition for the worst-case $\Delta\mu$ perturbation:

$$\underbrace{\frac{\partial\, E_t^{\eta}[dJ]}{\partial \Delta\mu}}_{\text{marginal utility impact}} + \lambda_t \underbrace{\frac{\partial\, R\left(\eta_t^{J_x}\right)}{\partial \Delta\mu}}_{\text{marginal entropy cost}} = 0.$$

The two terms in this equation capture the two opposing forces that determine the value of the perturbation. The left term captures the marginal harm to the investor's utility that results from increasing the perturbation, while the right term captures the associated increase in the detectability of the perturbed model. The worst-case model therefore assigns the largest perturbations to the parts of the reference model's dynamics that (I) *most damage utility* while (II) *remaining difficult to detect* statistically. Equations (6) and (7) show that (I) depends on the vector A of state-variable sensitivities. For instance, if $\Delta\mu$ is

[20] The difference between the exact and approximate expressions for the consumption-wealth ratio represents the error the investor makes in his optimal consumption-to-wealth choice. I evaluate the percentage consumption error induced by the approximation on a grid of values for the state variables, centered around their mean. I do this for the parameters used in the calibration (Section VI). I find that the errors are uniformly very small. For values of the state variables that are within one unconditional standard deviation of their mean, the errors are less than 0.05%. Even when all three persistent state variables are simultaneously three unconditional standard deviations from their mean, the percentage error is no greater than 0.60%. I also simulate the error in log consumption growth and find that the standard deviation of this error has an average across simulations of 0.034%, a very small value in comparison with the consumption moments.

the perturbation to the mean of the jump in $Y_{2,t}(i)$, and $\hat{A}_2(i) > 0$, then under the worst-case model $\Delta\mu < 0$, as this lowers the investor's utility.

The characteristics of the worst-case solution are nicely illustrated by the drift perturbation, for which an explicit solution is available. Let $\hat{A} = [\hat{A}_1', \hat{A}_2']'$ correspond to the partition $Y = [Y_1', Y_2']'$ and recall that the drift perturbation is given by $\Sigma_t h_t$, where $h_t = [0, h_{2,t}']'$. Appendix C shows that, under the worst-case model,

$$\Sigma_{2,t} h_{2,t} = -\frac{1}{\tilde{\lambda}} H_q \hat{A}_2 q_t^2, \tag{8}$$

where $\tilde{\lambda}$ is a constant that is part of the equilibrium solution for the λ_t process. The expression shows how the drift perturbations depend on the values of the components of $\hat{A}_2$. As pointed out above, the sign of a perturbation is opposite to that of the corresponding $\hat{A}_2$ loading. When $\hat{A}_2(i) > 0$, and the investor's utility is increasing in $Y_{2,t}(i)$, the worst-case perturbation *decreases* the drift in $dY_{2,t}(i)$. In addition, the size of the perturbation in $dY_{2,t}(i)$ is proportional to the magnitude of $\hat{A}_2(i)$, reflecting the sensitivity of the investor's utility to the state variable $Y_{2,t}(i)$. As the level of uncertainty fluctuates over time, so should the overall magnitude of the perturbations. This is captured by their dependence on q_t^2. Finally, the value of $\tilde{\lambda}$ controls the average size of the perturbations. This value depends inversely on φ, which controls the level of model uncertainty. Increasing φ increases uncertainty and decreases $\tilde{\lambda}$, resulting in larger perturbations, while decreasing φ increases $\tilde{\lambda}$ and shrinks the perturbations.

V. Asset Pricing

Since the representative investor makes decisions using the worst-case model, expected returns are determined by his Euler equation under the *worst-case* expectation. However, we are interested in expected returns under the *reference* model, since it is the best estimate of the data generating process based on historical data. To get reference-measure expected returns, one needs to adjust for the difference in expectations between the worst-case and reference probabilities. This difference represents an uncertainty premium that the investor is willing to pay to hedge himself against model uncertainty concerns.

A. *Pricing Kernel*

Since the standard Euler equation holds under the worst-case measure, the pricing kernel under that measure is the "usual" one, corresponding in this case to the Epstein and Zin (1989) pricing kernel. The expression for the log pricing kernel is then given by $d\ln M_t = -\theta\delta dt - \frac{\theta}{\psi} d\ln C_t - (1-\theta) d\ln R_{c,t}$, where $\theta = \frac{\gamma}{\rho}$ and $\frac{dR_{c,t}}{R_{c,t}} = \frac{dP_{c,t}+C_t dt}{P_{c,t}}$ is the instantaneous return on the aggregate consumption claim (aggregate wealth). As usual, when $\theta = 1$, this reduces to the (log) kernel under CRRA expected utility. The Internet Appendix shows that the

consumption-wealth ratio $(C_t/P_{c,t})$ equals $\delta \exp(-\rho g(\tilde{Y}))$, where $g(\tilde{Y}_t)$ is given by Proposition 1. Using Itô's lemma, one can then solve for $d\ln R_{c,t}$ in terms of dY_t. Substituting this and $d\ln C_t$ into $d\ln M_t$ gives the following expression for the pricing kernel under the worst-case measure:

$$d\ln M_t = -[\theta\delta + (1-\theta)\delta \exp(-\rho A_0 - \rho A'\tilde{Y}_t)]dt - \Lambda' dY_t, \tag{9}$$

where $\Lambda = (\frac{\theta}{\psi}\delta_c + (1-\theta)[\rho\hat{A} + (1-\rho)\delta_c])$. Λ is this kernel's vector of risk prices for the economy's shocks. In general, $\delta_c'\Lambda = 1-\gamma$, that is, the price of risk for the immediate consumption shock is the investor's RRA. When $\theta = 1$, the kernel reduces to that of power utility, and $\Lambda = (1-\gamma)\delta_c$, so all risk prices aside from that of the consumption shock's are zero. Note that, as (9) is the pricing kernel under the worst-case measure, it does not contain explicit uncertainty terms, but these will show up below due to an adjustment from the worst-case to the reference-model probabilities.[21]

A.1. The Risk-Free Rate

As the risk-free return is known with certainty at time t, it is identical under the worst-case and benchmark probabilities and no measure adjustment is necessary. Hence, it is given by $E_t^{\eta}[-\frac{dM_t}{M_t}]$, which implies

$$\begin{aligned} r_{f,t} = {} & \theta\delta + (1-\theta)\delta \exp(-\rho A_0 - \rho A'\tilde{Y}_t) + \Lambda^T(\mu(Y_t) + \Sigma_t h_t) \\ & - \frac{1}{2}\Lambda^T \Sigma_t \Sigma_t^T \Lambda - l_t^{\eta\prime}(\psi^{\eta}(-\Lambda) - 1). \end{aligned} \tag{10}$$

The explicit impact of model uncertainty on the risk-free rate is through the drift perturbation term $\Sigma_t h_t$ and the changes in the jump intensity l_t^{η} and the moment-generating function $\psi^{\eta}(-\Lambda)$. These perturbations decrease expected consumption growth and increase the precautionary savings motive by increasing expected future variance and the immediate probability and variance of a jump shock. Both effects lead to a lower equilibrium risk-free rate. Moreover, when q_t^2 rises and increases uncertainty, the perturbations are amplified and the impact on the risk-free rate increases.

B. Equity

A share of the stock market is modeled as a claim to the per-share dividend stream D_t. Let $v_{m,t}$ denote the log price–dividend ratio of the market and let $R_{m,t}$ denote the cumulative return through time t on a strategy that holds the market portfolio and fully reinvests all proceeds. Then $d\ln R_{m,t}$ is the instantaneous log market return. To solve for $v_{m,t}$, I use the

[21] Two points worth making are: (1) the pricing kernel under the reference measure is given by $d\ln M_t + d\ln \eta_t$, which explicitly accounts for the probability adjustment in the expression for the kernel, and (2) though the uncertainty terms do not show up explicitly in (9), uncertainty has an important impact in determining the vectors A and λ in (9).

Campbell and Shiller (1988) log-linearization of the instantaneous log return, adapted to continuous time as in Eraker and Shaliastovich (2008): $d\ln R_{m,t} = \kappa_{0,m}dt + \kappa_{1,m}dv_{m,t} - (1-\kappa_{1,m})v_{m,t}dt + d\ln D_t$, where $\kappa_{0,m}$ and $\kappa_{1,m}$ are the log-linearization constants. I conjecture that $v_{m,t}$ takes the following functional form:

$$v_{m,t} = A_{0,m} + A_m' Y_t. \tag{11}$$

Substituting for $dv_{m,t}$ gives the log market return in terms of dY_t:

$$d\ln R_{m,t} = \kappa_{0,m}dt - (1-\kappa_{1,m})(A_{0,m} + A_m' Y_t)dt + B_r' dY_t, \tag{12}$$

where $B_r = (\kappa_{1,m}A_m + \delta_d)$ is the vector of loadings on the shocks to the state vector. The return must satisfy the Euler equation under the worst-case measure. As the Internet Appendix shows, this implies a system of equations whose solution are the values of $A_{0,m}$ and A_m, verifying the conjectured $v_{m,t}$.

Combining equation (12) with equation (1) or (2) provides the equity return dynamics under the reference and worst-case models, respectively. Since Y_t is an affine jump-diffusion process under either model, as well as under the risk-neutral probability measure (as shown in the Internet Appendix), so is the (log-linearized) market return. Hence, in its reduced-form representation the model's market return is similar to the models used by widely cited (reduced-form) option pricing studies, such as Duffie, Pan, and Singleton (2000) and Broadie, Chernov, and Johannes (2007). The closest counterpart in Duffie, Pan, and Singleton (2000) is the SVJJ model, which Broadie, Chernov, and Johannes (2007) also use (they call it the SVCJ model). Both SVJJ and the model here feature a persistent square-root stochastic volatility process whose innovations are negatively correlated with returns. Both models also contain jumps in returns and volatility that occur simultaneously. Like the SVJJ model, the calibrated version of this paper's model (see Section VI) also features gamma-distributed volatility jumps.

The models also have important differences. The jump intensity here, l_t, is time-varying, while in SVJJ it is constant. In the SVJJ model, return jumps are normal conditional on the jump in volatility. In contrast, equation (12) shows that, owing to cross-equation restrictions, a jump in a component of Y_t (such as q_t) induces a proportional jump in the log return, and therefore both jumps have the same distribution (a gamma in the calibration). Moreover, the calibrated model contains a jump in equity returns that is not simultaneous with a jump in volatility, since it is induced by a jump in the rate of cash flow growth. Finally, it is important to keep in mind that the equilibrium model in this paper imposes tight cross-equation restrictions on the reduced-form parameters, whereas in reduced-form models they can be independently and freely chosen.

C. The Equity Premium

Having solved for A_m, one can solve for the equity premium. The Euler equation under the worst-case measure implies $E_t^\eta\left[d(M_t R_{m,t})\right] = 0$ (equivalently, $M_t R_{m,t}$ is an η-martingale). Applying Ito's lemma with jumps and substituting in $r_{f,t} = -E_t^\eta[-\frac{dM_t}{M_t}]$ gives an expression for $\frac{1}{dt}E_t^\eta[\frac{dR_{m,t}}{R_{m,t}}] - r_{f,t}$, the equity premium under the worst-case probabilities. The equity premium that one would observe under the reference-model probabilities is then obtained by adding the adjustment $E_t[dR_{m,t}/R_{m,t}] - E_t^\eta[dR_{m,t}/R_{m,t}]$. The resulting expression for the equity premium is

$$\frac{1}{dt}E_t\left[\frac{dR_{m,t}}{R_{m,t}}\right] - r_{f,t} = B_r'\Sigma_t\Sigma_t^T\Lambda - B_r'\Sigma_t h_t + l_t'(\psi(B_r) - 1)$$
$$- (l_t^\eta \cdot \psi^\eta(-\Lambda))'\left(\frac{\psi^\eta(-\Lambda + B_r)}{\psi^\eta(-\Lambda)} - 1\right), \tag{13}$$

where division in the last term is componentwise. The first term on the right side is (negative) the continuous covariance of the market return and the pricing kernel. If there are no jumps or model uncertainty, this term is the entire equity premium. The next term is due to the drift perturbation. It increases the equity premium when the components of B_r and $\Sigma_t h_t$ have opposite signs, which is normally the case. This is because, if the investor dislikes a state variable, the corresponding perturbation component is positive and the market return typically responds negatively to an increase in that variable (i.e., $B_r(i) < 0$).[22] The second line of (13) can be viewed as the jump return premium. Uncertainty increases this term through its impact on the jump intensity l_t^η and the moment-generating function ψ^η (as indicated by their dependence on η). It amplifies the investor's assessment of the frequency l_t^η of jumps and changes their distribution (via ψ^η) by increasing the likelihood of bigger, more negative jumps. Both effects increase the jump risk premia. When q_t^2 rises and increases uncertainty, the perturbations become larger and the return premium emanating from the diffusion and jump-related uncertainty terms increases. Finally, since uncertainty impacts the equilibrium values of B_r and Λ, it also has an important indirect influence on (13).

D. The Variance Premium

Let $* \in \{P, \eta, Q\}$ indicate the physical/reference (P), worst-case (η), or risk-neutral (Q) probability measures. The $*$-expectation of integrated return variance from time t to $t+1$ is $E_t^*[\int_t^{t+1}(d\ln R_{m,s})^2]$, or equivalently $\int_t^{t+1} E_t^*(d\ln R_{m,s})^2$, and the variance premium is the difference in this quantity between $* = Q$ and $* = P$. I explain here how uncertainty is involved in generating a variance premium, leaving complete analytical expressions to the

[22] This relationship holds when $\psi > 1$ (which is used in the calibration), but can be reversed when $\psi < 1$.

Internet Appendix. Consider the integrand above for the first instant, that is, for $s = t$. By equation (12) it is given by

$$E_t^*(d\ln R_{m,t})^2 = B_r' \Sigma_t \Sigma_t' B_r + B_r^{2'} \left[E^*\left(\xi_t^2\right) \cdot l_t^*\right],$$

where $l_t^* = l_1^* q_t^2$ is the jump intensity under the measure $*$ (recall that $l_1^P = l_1$, $l_1^\eta = \exp(a) l_t$, while l_1^Q is derived in the Internet Appendix). Note that the jump component of variance depends on the probability measure, but the diffusion component does not. Due to his model uncertainty, the investor is worried about underestimating the frequency and magnitude of jumps, which are therefore amplified under the worst-case model. This implies that

$$E_t^\eta(d\ln R_{m,t})^2 - E_t^P(d\ln R_{m,t})^2 = B_r^{2'}\left[E^\eta\left(\xi_t^2\right) \cdot l_1^\eta - E^P\left(\xi_t^2\right) \cdot l_1^P\right] q_t^2 > 0. \tag{14}$$

Reflecting the concerns about jumps, the expected variance over t to $t + dt$ is higher under the worst-case model. Note also that the size of the difference across the two models is a multiple of q_t^2, so it covaries with the level of uncertainty.

The risk-neutral probabilities, which are determined by (9) under the worst-case measure, tilts the probabilities of the worst-case measure towards states in which marginal utility is high, in particular, states with large negative jump realizations. This augments the gap in expectations from (14) and gives

$$E_t^Q(d\ln R_{m,t})^2 - E_t^P(d\ln R_{m,t})^2 = B_r^{2'}\left[E^Q\left(\xi_t^2\right) \cdot l_1^Q - E^P\left(\xi_t^2\right) \cdot l_1^P\right] q_t^2 > 0. \tag{15}$$

To compare the integrated expected variance across the different probabilities, one still needs to consider how the quantity $E_t^*(d\ln R_{m,u})^2$ evolves as u increases. In fact, this quantity will drift higher under η relative to P. A reason for this is the drift perturbation (8), which increases the drift of state variables that decrease the investor's utility, including the variables that drive volatility. Combined with (14), this gives $\int_t^{t+1} E_t^\eta(d\ln R_{m,s})^2 - \int_t^{t+1} E_t^P(d\ln R_{m,s})^2 > 0$. Since the size of the drift perturbation (equation (8)) covaries with q_t^2, so does the difference in integrated expected variance. The risk-neutral measure further augments the difference in drifts and therefore $vp_t = \int_t^{t+1} E_t^Q(d\ln R_{m,s})^2 - \int_t^{t+1} E_t^\eta(d\ln R_{m,s})^2 > 0$.

Two further points should be noted. First, vp_t is a good filter for q_t^2 and the level of uncertainty. The reason is that differencing the variance expectations filters out variance due to shocks that have little or no impact on utility and are therefore unaffected by the worst-case model. This leaves the parts that are amplified by uncertainty, making vp_t a good filter for the level of uncertainty. Second, since the level of uncertainty is a factor driving variation in the equity premium, vp_t will predict excess equity returns with a positive predictive coefficient. Moreover, its "filtered" nature implies that it should be a better predictor of excess returns than expectations of overall variance.

VI. Calibration

I now parameterize and calibrate a version of the model framework with uncertainty to quantitatively match a broad set of moments of cash flows, returns, and option prices. The results illustrate quantitatively the impact of model uncertainty concerns on asset prices.

A. Reference Model Specification

The reference model is an extension of the long-run risks model in Bansal and Yaron (2004) (BY) and Bansal, Kiku, and Yaron (2007). The dynamics of this model fit the cash flow data well and also provide multiple channels on which model uncertainty can act. As in BY, there is a small but persistent component in consumption and dividend growth, denoted by x_t. The cash flow processes are given by

$$d\ln C_t = \left(\mu_c + x_t - \frac{1}{2}\Phi_c^2\sigma_t^2\right)dt + \sigma_t\Phi_c dZ_{c,t}$$

$$d\ln D_t = \left(\mu_d + \phi x_t - \frac{1}{2}\Phi_d^2\sigma_t^2\right)dt + \sigma_t\Phi_d dZ_{d,t}.$$

The parameter ϕ represents the sensitivity of dividend growth to x_t and is greater than one, reflecting the greater volatility of dividends relative to consumption. I assume there is no ambiguity about the structure of these immediate cash flow growth rates and let $Y_{1,t} = (\ln C_t, \ln D_t)'$. The conditional variance of the consumption and dividend growth streams is driven by the stochastic process σ_t^2, which I assume follows an autoregressive process. I also let q_t^2 follow an autoregressive process. I assume there is uncertainty about the dynamics of the three persistent state variables and therefore let $Y_{2,t} = (\sigma_t^2, x_t, q_t^2)$.

In summary, the state vector Y_t and transition matrix $\mathcal{K}$ are given by

$$Y_t = \begin{pmatrix} \ln C_t \\ \ln D_t \\ \sigma_t^2 \\ x_t \\ q_t^2 \end{pmatrix}, \qquad \mathcal{K} = \begin{pmatrix} 0 & 0 & -\frac{1}{2}\Phi_c^2 & 1 & 0 \\ 0 & 0 & -\frac{1}{2}\Phi_d^2 & \phi & 0 \\ 0 & 0 & \rho_\sigma & 0 & 0 \\ 0 & 0 & 0 & \rho_x & 0 \\ 0 & 0 & 0 & 0 & \rho_q \end{pmatrix},$$

I fix the value of the drift vector μ by setting $E(d\ln C_t) = \mu_c$, $E(d\ln D_t) = \mu_d$, $E(x_t) = 0$, $E(\sigma_t^2) = 1$ and $E(q_t^2) = 1$, which is without loss of generality. Finally, the diffusion covariance matrix and jumps specification are

$$\Sigma(Y_t)\Sigma(Y_t)' = \begin{bmatrix} H_\sigma\sigma_t^2 & 0 \\ 0 & H_q q_t^2 \end{bmatrix}, \quad l_t = l_1 q_t^2, \quad \xi_x \sim N(0, \sigma_x^2), \quad \xi_q \sim \Gamma\left(\nu_q, \frac{\mu_q}{\nu_q}\right),$$

where $H_\sigma = \text{diag}(\Phi_c^2,\ \Phi_d^2)$ and $H_q = \text{diag}(\Phi_\sigma^2,\ \Phi_x^2,\ \Phi_q^2)$. Hence, the diffusions are uncorrelated. Note that there are jumps in both x_t and q_t^2, denoted ξ_t and ξ_q,

respectively. The mean jump intensity is given by $l_1 = (0, 0, 0, l_{1,x}, l_{1,q})'$ and the jump size distributions are shown above. The jumps in x_t have a zero-mean normal distribution, while the jump sizes in q_t^2 have a gamma distribution. Specifying a gamma jump size guarantees that q_t^2 remains positive. The parameter ν_q controls the "shape" of the gamma distribution, while μ_q controls its "scale." This parameterization implies that $E[\xi_q] = \mu_q$. For $\nu_q = 1$, the value used in the calibration, the gamma distribution reduces to an exponential distribution.

There are several motivations for the choice of two volatility processes and the partition of Y_t. One reason is to separate variation in pure cash flow volatility from time-varying model uncertainty. In most structural models, the majority of return volatility comes from cash flow volatility, as is also the case here. However, it need not be the case that uncertainty move in lock-step with cash flow and return volatility, and creating separate volatility processes allows the model to capture this potential separation. For the partitioning of Y_t, it is reasonable that model uncertainty should be much less important for immediate cash flows than for the dynamics of the persistent variables. The immediate cash flow growth rates are comparatively easy to measure and have low persistence, in contrast to the state variable dynamics, which are hard to estimate and potentially quite persistent. Note though, that uncertainty about the dynamics of dY_2 directly affects future expectations of dY_1, so there definitely *is* uncertainty about future cash flow growth rate dynamics.

B. Parameter Values

I calibrate the model using the following guidelines. I aim to find parameter values for the model specification such that (i) the reference model's time-averaged consumption and dividend growth statistics are consistent with salient features of the consumption and dividend data; (ii) the model generates unconditional moments of asset prices, such as the equity premium and the risk-free rate, that match those in the data; and (iii) the model matches moments of market return volatility, the VIX, and the variance premium, as well as the projections of excess stock returns on the variance premium. Finally, the calibration also compares the model-generated implied volatility curves for 1-, 3-, and 12-month maturities with their empirical counterparts. Table IV reports the parameter values for the calibration, normalized to a monthly interval, that is, $\Delta t = 1$ is one month.

The cash flow parameters are similar to those in the long-run risk models of Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2007), though the expected growth component x_t is somewhat less persistent.[23] As Table IV shows, the volatility and uncertainty processes, σ_t^2 and q_t^2, are also persistent, though significantly less than the volatility process in BY. Turning to the jump

[23] In comparing the parameter values to those of a discrete-time model, it is important to remember that ρ_x in this model's continuous-time formulation maps to $\exp(\rho_x)$ in a discrete-time setup.

Table IV
Calibration: Model Parameters

This table presents the parameters for the reference model used in the model calibration (see Section VI.A). The calibration results are presented in Tables V and VI. Parameter values are normalized for a monthly interval, that is, $\Delta t = 1$ is one month.

Preferences	δ	RRA	ψ	φ	
	$-\ln 0.999$	5	2.0	0.0048	
Consumption	$E[\Delta c]$	Φ_c			
	0.0016	0.0066			
Persistent mean component	ρ_x	Φ_x	$l_1(x)$	σ_x	
	-0.025	$0.042 \times \Phi_c$	1.0/12	$2.25 \times \Phi_x$	
Dividends	$E[\Delta d]$	ϕ	Φ_d		
	0.0016	3	$6.0 \times \Phi_c$		
Cash-flow variance	ρ_σ	Φ_σ			
	-0.1	0.30			
Uncertainty	ρ_q	Φ_q	$l_1(q)$	μ_q	ν_q
	-0.2238	0.25	0.75/12	1.5	1

parameters, jumps in x_t have a standard deviation that is 2.25 times the average volatility of the x_t diffusion, and occur at an average rate of one jump per year. In contrast to the rare-disasters literature, these jumps are infrequent but not rare, and are (potentially) large compared to the diffusion but not "disastrous." Jumps in q_t^2 occur at an average rate of 0.75 jumps per year with a mean jump size of 1.5. These jumps generate spikes in the level of uncertainty, which is important for capturing the large variance premium and options prices. The parameter values are chosen to quantitatively match the high skewness and kurtosis in the dynamics of conditional variance. They also help capture the negative skewness and excess kurtosis of equity returns. The frequency used for these jumps is consistent with findings from studies on the time-series properties of returns and option prices (see, for example, the review in Singleton (2006)).

Table IV also shows the preference parameters. The relative risk aversion of the representative investor is set to five. This is right in the middle of the range considered plausible by Mehra and Prescott (1985) and is far lower than the levels of risk aversion needed to match the equity premium in habits or long-run risk models. The value of the IES is set to two, which corresponds to the value used in Bansal, Kiku, and Yaron (2007). An IES above one helps dampen the volatility of the risk-free rate. There is a long-standing debate in the literature about the value of the IES (see Beeler and Campbell (2012) and Bansal, Kiku, and Yaron (2012)).

To set φ, which controls the level of model uncertainty, I follow Anderson, Hansen, and Sargent (2003) and calculate the *detection-error* probability for the resulting worst-case model. This is the probability that an investor would *incorrectly* reject the worst-case model in favor of the reference model using a

likelihood ratio test. If the detection-error probability is reasonably high, then it is difficult to reject the worst-case model and the investor has reason to be concerned that it may in fact be the true model. Two essentially indistinguishable models will have a 50% detection error probability as they are equally likely to fit a given data sample. Anderson, Hansen, and Sargent (2003) argue that 10% is a reasonable bound for the detection-error probability of the worst-case model, in line with the conventional 10% Type-1 error probabilities used in statistics. Given η_t for the worst-case model, the detection-error probability can be calculated precisely as explained in the Internet Appendix.[24] Besides φ (and η_t), the detection-error probability also depends on the length of data history the investor can use for inference. This history is limited by both the availability of data records and also by any structural breaks that may have occurred within the sample. The φ used in the calibration corresponds to about 11% detection-error probability for the post-war sample. The same φ implies detection-error probabilities of 47.5%, 26.0%, and 16.4% for sample lengths of 1, 20, and 40 years, respectively. Note that the decay in the probability is slow, highlighting the difficulty in distinguishing the worst-case and reference models and the investor's enduring concern about using the correct model.

C. Results

Tables V and VI provide the empirical moments and the corresponding statistics for the calibrated model. To assess the model fit to the data, I provide model-based finite sample statistics. Specifically, I present the model-based 5%, 50%, and 95% percentiles for the statistics of interest generated from 10,000 simulations, each based on the same sample length as its data counterpart. The time increment used in the simulations is one month. For all data, I report moments for both the pre-crisis sample, which ends in June 2007, and the full sample, which ends in December 2009. For the consumption and dividend dynamics, I use the longest sample available, which starts in 1930. Hence, the full-sample simulations are based on 960 monthly observations that are time-averaged to an annual sample of length 80, as in the data. I provide similar statistics for the the mean and volatility of the market return, risk-free rate, and price–dividend ratio. For the variance premium-related statistics the data are monthly and available starting in 1990. Thus, the model's variance premium-related statistics are based on the last 240 monthly observations in each of the 10,000 simulations. The model is calibrated to match the simulation medians with the pre-crisis estimates rather than the full sample's. For the cash flow and equity return moments, this choice makes no practical difference, as indicated below. Where it does make a difference is for some second and higher moments of

[24] For simple i.i.d. environments, there is an analytical formula while for the calibrated model this calculation requires solving a set of ODEs for the characteristic function of η_t and then calculating a Fourier inversion numerically. To my knowledge, this is the first paper that reports exact detection error probabilities for a jump-diffusion setting. Maenhout (2006) is the first to describe how this can be done using the Fourier method.

the variance dynamics. These moments, calculated from the post-1990 sample, were impacted greatly by the huge spikes in volatility that occurred during the crisis, particularly in October and November of 2008. To avoid matching the simulation medians to these crisis-sensitive moments of the shorter sample, I calibrate the model to the pre-crisis sample. Nevertheless, it is important to note that the calibrated model does allow for such outlying spikes. This is reflected in the calibration's finite-sample percentiles and is discussed below. Reporting both the pre-crisis and full-sample data estimates provides a fuller picture of both the data and the model fit and also allows for an interesting comparison of the finite-sample properties of the model with those implied by the data.

It is important to note that the dynamics being simulated are those of the *reference* model. These are the right dynamics to use for reporting simulation moments under the view that the calibration's reference model is the one used by investors and that it provides a good fit to the historical data. It also disciplines the choice of simulated dynamics, with asset prices reflecting the impact of model uncertainty. Under this view, the data point estimates should be within the 5% to 95% percentile intervals generated by the model simulations. For completeness, I also provide Newey–West heteroskedasticity and autocorrelation consistent (HAC) robust standard errors for the data statistics.

C.1. Cash Flows and Equity Returns

The top panel in Table V shows that the reference model captures well several key moments of annualized consumption and dividend growth. The data-based mean and volatility of consumption and dividends growth are close to the median estimates from the model and fall within the 90% confidence interval. Moreover, the distribution of model moments is very reasonable, even in the presence of jumps. The autocorrelations in the cash flow processes are also close to their model counterparts. Hence, the calibrated reference model does a good job matching the cash flow data and is a quite reasonable specification for investors to use as their reference model. Comparing the pre-crisis and full-sample estimates shows that they do not differ much and that the model fit is very good for both.

The bottom panel of Table V presents the model implications for equity returns. It shows annual data on the market return, risk free rate, and price–dividend ratio. The model does a good job capturing the large equity premium and volatility of excess returns. This is the case even though the representative investor's relative risk aversion is only five. At the same time, the model matches the high volatility of the market return. The table further shows that the model's risk-free rate has a low mean and volatility, similar to the data, though the model's mean is a bit high and the volatility a bit low.[25] The rows

[25] A related issue is the debate over the value of the IES, which has been estimated in the literature by regressing consumption growth on the lagged risk-free rate. Estimating this regres-

Table V
Model Calibration Results

This table presents moments of consumption/dividend dynamics and asset prices from the data and for the model of Table IV. The data are real and sampled at an annual frequency, except for the lines labeled with "(M)," which are sampled at a monthly frequency. Data estimates are reported for samples starting in 1930 and ending in 2007 and 2009, as indicated. Standard errors are given in parentheses and are Newey–West with four lags. For the model, I report percentiles of these statistics based on 10,000 model simulations with each statistic calculated using a sample size equal to its 1930 to 2009 data counterpart. E is the average value, σ the standard deviation, $AC1$ the autocorrelation, $skew$ the skewness, and $kurt$ the kurtosis for the indicated series. Δc and Δd are the log growth rates in annual aggregate consumption and per-share dividends, respectively. In both the data and the model, consumption and dividends are time-averaged to an annual frequency. $r_m - r_f$ denotes the excess log return of the stock market over the three-month Treasury bill. $p - d$ denotes the log price–dividend ratio for the stock market with dividends adjusted for repurchases.

	Data				Model		
	2007.6		2009.12		5%	50%	95%
Cash Flow Dynamics							
$E[\Delta c]$	1.88	(0.32)	1.81	(0.32)	0.98	1.92	2.86
$\sigma(\Delta c)$	2.21	(0.52)	2.21	(0.50)	2.03	2.45	2.99
$AC1(\Delta c)$	0.43	(0.12)	0.44	(0.12)	0.26	0.46	0.63
$E[\Delta d]$	1.54	(1.53)	0.61	(1.67)	−1.79	1.83	5.50
$\sigma(\Delta d)$	13.69	(1.91)	15.29	(2.24)	10.41	12.34	14.60
$AC1(\Delta d)$	0.14	(0.14)	0.20	(0.13)	0.12	0.31	0.48
$corr(\Delta c, \Delta d)$	0.59	(0.11)	0.60	(0.09)	0.00	0.24	0.46
Returns							
$E[r_m - r_f]$	5.41	(2.09)	4.93	(2.08)	3.16	6.29	9.50
$E[r_f]$	0.82	(0.35)	0.79	(0.34)	1.08	1.52	1.94
$\sigma(r_m - r_f)$	19.48	(2.35)	20.15	(2.24)	15.58	18.02	21.03
$\sigma(r_f)$	1.89	(0.17)	1.89	(0.16)	0.65	0.88	1.23
$E[p - d]$	3.15	(0.07)	3.14	(0.06)	2.76	2.83	2.90
$\sigma(p - d)$	0.31	(0.02)	0.31	(0.02)	0.13	0.16	0.21
$skew(r_m - r_f)$ (M)	−0.43	(0.54)	−0.46	(0.52)	−0.86	−0.27	0.10
$kurt(r_m - r_f)$ (M)	9.93	(1.26)	9.63	(1.23)	4.01	5.73	10.80

labeled "skew" and "kurt" give the skewness and kurtosis of *monthly* excess returns for the sample (1930 to 2009).[26] Since the higher moments of the stock return distribution have an important effect on option prices and the implied volatility skew, it is important that the model's dynamics be consistent with them while attempting to match option prices. The table shows that is indeed the case. The one moment for which the model falls somewhat short is in gen-

sion coefficient on quarterly data within the model simulations, I find that the 5%, 50%, and 95% values for the coefficient are 0.71, 1.39, and 2.03. Though these values are higher than the estimate reported by Beeler and Campbell (2012), it is clear that the IES estimate is biased downwards significantly. Moreover, as Bansal, Kiku, and Yaron (2012) point out, adding realistic measurement error to the real risk-free rate will both bias the estimate down further and increase its sampling variability.

[26] For comparison, the post-war (1949 to 2009) estimates for the monthly excess return kurtosis and skewness are 5.97 (1.48) and –0.84 (0.33), respectively.

Table VI
Variance Premium

This table presents moments from the data and for the model of Table IV for conditional variance and the variance premium (top panel) and regressions of excess stock market returns on the variance premium (bottom panel). The data estimates are reported for the pre-crisis sample, 1990.1 to 2007.6, and the full sample, 1990.1 to 2009.12, as indicated, and correspond to a monthly frequency. For the model, I report percentiles of these statistics based on 10,000 simulations, with each statistic calculated using a sample size equal to its full-sample data counterpart. Standard errors are given in parentheses and are Newey–West with four lags. *VP* is the variance premium, $\text{var}_t^P(r_m)$ is the physical (i.e., true) forecast of return variance, and $\text{var}_t^Q(r_m)$ is the risk-neutral expectation of variance, which is given by VIX^2. $\beta(i)$ and $R^2(i)$ denote the slope coefficient and R^2 for predictive regressions of excess stock market returns over horizons of $i = 1, 3,$ and 6 months on the lagged variance premium. For the regressions, excess returns are expressed as annualized percentages.

	Data				Model		
	2007.6		2009.12		5%	50%	95%
Variance Premium							
$E[VP]$	11.05	(1.07)	10.55	(2.62)	5.75	8.17	13.01
$\sigma(VP)$	7.63	(1.09)	8.47	(1.09)	3.40	7.53	16.31
$skew(VP)$	2.43	(0.60)	2.19	(0.38)	1.54	2.69	4.31
$kurt(VP)$	12.34	(3.52)	9.67	(2.09)	5.67	11.83	26.33
$AC1(\text{var}_t^P(r_m))$	0.82	(0.04)	0.81	(0.05)	0.76	0.85	0.92
$AC1(\text{var}_t^Q(r_m))$	0.79	(0.05)	0.82	(0.05)	0.75	0.84	0.92
$\sigma(\text{var}_t^P(r_m))$	17.41	(2.21)	31.00	(8.36)	8.83	14.97	29.80
$\sigma(\text{var}_t^Q(r_m))$	23.97	(3.15)	36.58	(8.07)	11.36	21.69	45.26
$skew(\text{var}_t^Q(r_m))$	2.02	(0.49)	3.38	(0.51)	0.88	2.19	3.75
$kurt(\text{var}_t^Q(r_m))$	8.98	(2.26)	19.21	(5.07)	3.96	9.18	21.23
Predictability							
$\beta(1)$	0.74	(0.34)	0.80	(0.36)	−0.09	1.10	2.65
$R^2(1)$	1.44	(1.48)	1.72	(1.64)	0.03	1.74	7.32
$\beta(3)$	0.87	(0.25)	0.99	(0.25)	−0.14	0.94	2.22
$R^2(3)$	6.15	(4.50)	6.93	(4.53)	0.06	3.91	16.93
$\beta(6)$	0.51	(0.22)	0.76	(0.23)	−0.22	0.75	1.78
$R^2(6)$	4.47	(4.77)	7.28	(5.33)	0.07	5.31	24.53

erating a large enough volatility for the price–dividend ratio. Comparing again the pre-crisis and full-sample estimates, the table shows there is not much difference.

C.2. Variance Dynamics and Premium

Table VI provides statistics for integrated variance and the variance premium, all at the monthly horizon. I focus first on a comparison of the model with the pre-crisis estimates and then discuss the full-sample moments. Note first that the model successfully generates a large average variance premium ($E[VP]$), comparable to that in the data. Hence, as in the data, index options in the model embed a high insurance premium. It is important to note that this

large premium is by no means implied by the fact that the model matches the equity premium or the level of return volatility. To illustrate this, I solve for the variance premium implied by the calibrated model in Bansal, Kiku, and Yaron (2007), a leading asset pricing model that successfully matches the equity premium and high equity return volatility. I find that it is approximately 0.35, which is far smaller than both the data and the values generated by the model of this paper.

Table VI also shows that the model captures the high volatility of the conditional variance premium. This reflects the impact of time-variation in uncertainty concerns, as discussed in Section V. The table further shows that the model matches the high skewness and kurtosis of the variance premium, producing large values that are in line with the data. This is a result of spikes in the variance premium that arise due to jumps up in q_t^2 and the level of uncertainty.

There is a tension between generating a large variance premium and remaining consistent with the dynamics of conditional variance, which creates a challenge for models that confront both aspects of the data. The rest of the top panel in Table VI demonstrates that the model respects the properties of the two conditional variance series, $\text{var}_t^P(r_m)$ and $\text{var}_t^Q(r_m)$ (i.e., VIX^2). It shows that the empirical autocorrelation estimates for these series are close to the medians of the simulated series, and easily within the 90% confidence intervals. Conditional variance is persistent, but not extremely so. The next two lines in the table show that the model is in line with the high volatility in the conditional variances themselves and matches the relatively higher volatility in $\sigma(\text{var}_t^Q(r_m))$ compared to $\sigma(\text{var}_t^P(r_m))$. The final two lines of the table show that the squared VIX time series displays pronounced nonnormality, with a large positive skewness and kurtosis. The model matches these features well.

Section II highlights the predictive ability of the variance premium for excess equity returns in the data. In the model, the variance premium is a good filter for detecting fluctuations in the level of uncertainty. Since increases in uncertainty cause equity prices to fall and the expected return to rise, the variance premium predicts excess equity returns within the model. The bottom panel of Table VI shows that the model also captures this predictability quantitatively. It presents the projection coefficients in the data and in the model for predictive regressions of log excess returns on the variance premium for horizons of one, three, and six months. As the table shows, the projection coefficients have the right sign and the median values are roughly in line with the data estimates. As in the data, the model-based R^2s are quite large for these short horizons. The model median R^2 for the one-month-ahead projection is roughly 1.75% and the 90% finite sample distribution of R^2 clearly includes the 1.44% R^2 from the data. For the three- and six-month-ahead projections, the median R^2 increases to approximately 3.9% and 5.3% inside the model, which, though high, is similar to the 6.2% and 4.5% values in the data. Finally, in the model the variance premium is superior to the conditional variance as a predictor of equity returns, as first documented empirically by Bollerslev, Tauchen, and Zhou (2009). This is due to the fact that, while both σ_t^2 and q_t^2 drive variation

in the conditional variance, variation in the equity risk premium comes mostly from fluctuations in q_t^2, which the variance premium captures cleanly.

C.3. Pre-Crisis and Full Samples

Comparing the pre-crisis and full-sample estimates, there is a striking increase in the volatility and higher moments of the conditional variance series. A large portion of this increase is due to October and November of 2008. Excluding those two months, the estimates for $\sigma(\text{var}_t^P(r_m))$ and $\sigma(\text{var}_t^Q(r_m))$ are 22.6 and 29.5, respectively, while the skewness and kurtosis of $\text{var}_t^Q(r_m)$ are 2.22 and 9.25, respectively. The changes in the other statistics in Table VI are less dramatic. The mean variance premium decreases from about 11 to 10.5 and the volatility of the variance premium increases from 7.6 to around 8.1. The very large realized variance in October and November of 2008 generated a big positive payoff to the variance swap replication strategy discussed in Section II, reducing the estimate of the average premium paid by investors for this strategy. This reduction in the mean of VP was attenuated by the large variance premium that prevailed in the following months, when uncertainty was very high.

Note that the potential for a large change in the conditional variance moments across the two samples *is* consistent with the calibration's simulation percentiles. The percentiles indicate that, for the given sample length, these particular moments have a wide sampling distribution under the model due to the potential for large jump shocks. Notice that the full-sample estimates for the volatility and higher moments of the conditional variances are near the 95% of their model-based distributions. This is consistent with the crisis's volatility dynamics representing a relative outlier for this sample length. It also highlights the ability of the model to capture the possibility of this large change in moments across the samples, providing evidence in favor of the model dynamics. In contrast, the changes in the estimates are *not* consistent with the robust (HAC) standard errors for the pre-crisis sample, as most of the observed changes in the estimates are by more than four standard errors. Compared to the simulation percentiles, the robust standard errors are generally consistent for the cash flow and return moments, but greatly underestimate the width of the sampling distributions for the variance dynamics. This observation is itself also consistent with the model, as shown in Table VII, which reports HAC standard errors for the moment estimates computed within the model simulations. The table shows that HAC standard errors for the variance moments computed within the model are consistent with their empirical counterparts and tend to underestimate the sampling distribution.

C.4. Option Prices and the Volatility Surface

Figure 2 displays plots comparing the empirical and model-based implied volatility curves for 1, 3, and 12-month maturities. These provide an

Table VII
Calibration Results: Simulated HAC Standard Errors

This table presents the median HAC standard error corresponding to the given set of Monte Carlo (model-based) moments from Table VI. The sample of HAC standard errors is calculated from the same samples used to generate the model-based statistics in Table VI. As with the data-based standard errors, which are given in the column labeled "Data," the model-based HAC standard errors are calculated using the Newey–West variance–covariance estimator with four lags. Symbols are defined in Table VI.

	Data 2007.6	Model 50%
Variance		
$E[VP]$	1.07	0.94
$\sigma(VP)$	1.09	1.40
$skew(VP)$	0.60	1.60
$kurt(VP)$	3.52	7.59
$\sigma(\text{var}_t^P(r_m))$	2.25	2.15
$\sigma(\text{var}_t^Q(r_m))$	3.14	3.54
$skew(\text{var}_t^Q(r_m))$	0.49	1.29
$kurt(\text{var}_t^Q(r_m))$	2.26	5.56

assessment of the model's ability to fit the cross-sectional patterns in implied volatilities.[27] To produce the model-based curves, I determine option prices for the calibrated model and solve for their Black–Scholes implied volatilities. As in the data, the options I price within the model are written on the equity price index excluding dividend payments (as opposed to a total return price index, which would include dividends). Option prices are calculated by first solving a system of ODEs for the characteristic function of the state vector under the Q-measure, as in Duffie, Pan, and Singleton (2000). I next use this to compute the characteristic function of the price index and calculate and invert the option transform based on the method of Carr and Madan (1999). The plots present the model and empirical implied volatilities for strikes ranging in moneyness (Strike/Spot Price) from 0.75 to 1.25. This represents a very wide range of strikes for short-maturity options. The empirical implied volatilities are the average daily implied volatilities for options on the S&P 500 index traded in the over-the-counter market. The data are obtained from Citigroup and cover October 1999 to June 2008.[28] For the calculation of the model-based option prices, the state vector Y_t is set equal to its unconditional mean.

[27] Note that fitting the implied volatility surface does *not*, by itself, imply that a model captures the variance premium. A model may violate aspects of the statistical (i.e., P-measure) distribution of returns while fitting the implied volatility surface. Matching the variance premium is a challenging test for equilibrium models because it requires directly capturing the *difference* between the P- and Q-measure distributions, that is, the "premium."

[28] The range of strikes available in the over-the-counter market for index options is often much broader than for exchange-traded options. See Foresi and Wu (2005) for more details on the over-the-counter market.

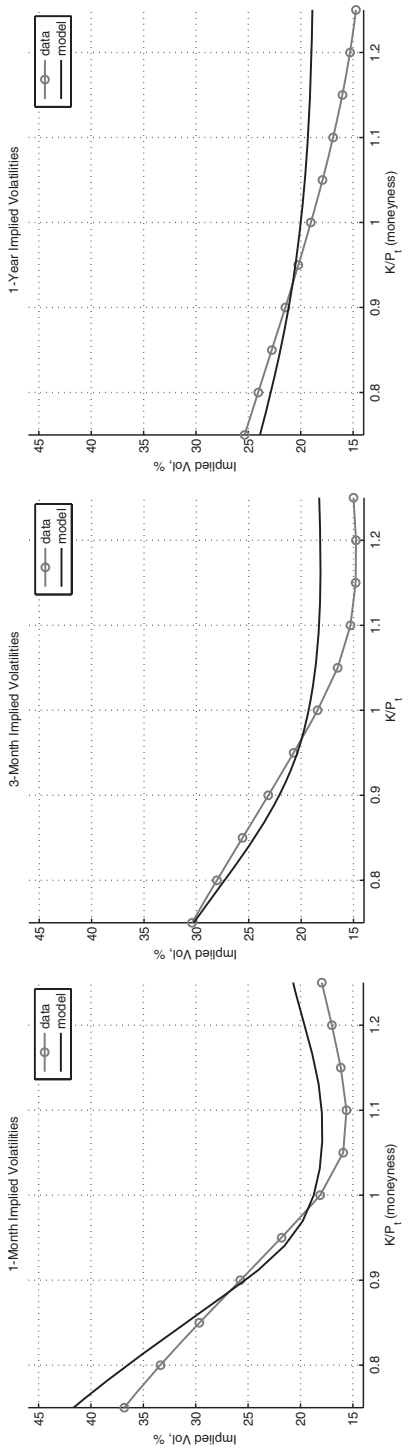

Figure 2. One-, 3-, and 12-month implied volatilities: Model and data. The figure plots comparisons of empirical and model-based implied volatilities for 1-, 3-, and 12-month maturities (left, center, and right plots, respectively) for the model of Table IV. Strikes are expressed in moneyness (Strike Price/Spot Price). The empirical curves are means of daily implied volatilities for S&P 500 index options for the period 1999.10 to 2008.6. The model-based curves are calculated for option prices obtained when the model's state vector is set equal to its unconditional mean.

The plots in Figure 2 show that the model does a good job matching both the overall shape and the level of the implied volatility curves. In particular, it captures the skew, the steep slope in implied volatilities for otm puts. Model uncertainty is instrumental in capturing this by increasing fears about the likelihood and magnitude of adverse jump shocks, which result in negative returns and generate a Q-measure return density with a heavy left tail. The implied volatility skew exists at all three maturities, but is particularly pronounced for the one-month maturity. The model captures the skew well at each of the three maturities, and does a very nice job replicating the decay in the skew as the horizon increases. As can be seen by comparing across the plots, the model further matches the shallow positive slope in the at-the-money implied vols, whereby the 12-month at-the-money vol is the highest and the 1-month vol the lowest. This is also clear from Figure IA.1 in the Internet Appendix, which displays separate plots of the empirical and model-based curves, further highlighting the similarities in their appearance.

In addition to doing a good job matching the very steep slope in the one-month implied volatility curve at low moneyness strikes, the model also replicates the smirk, the slight increase in implied volatility at high moneyness strikes. However, it is apparent that the model-based curve does not dip as low as the empirical one. The same thing is true for the 3- and 12-month plots—the model fit is very good for low moneyness strikes but not as good for high-moneyness strikes.

The Internet Appendix presents the results of a further exercise that use the data on one-month implied volatilities. It displays the time series for q_t^2 and σ_t^2 extracted by fitting the calibrated model each month to a set of one-month implied volatilities at different levels of moneyness. As reported there, the sample moments of the extracted time series are consistent with the moments implied by the model calibration.

D. The Impact of Uncertainty

As Section IV.B explains, the directions and magnitudes of the worst-case model perturbations depend on the equilibrium $\hat{A}$ coefficients, which encode the sensitivity of the value function to perturbations. Under the worst-case model, cash flow growth is lower, the drift and persistence in volatility is increased, and the jump intensities and magnitudes are amplified. The overall impact of uncertainty on equilibrium asset prices depends on the combined interacted effects of the worst-case perturbations. The comparative statics exercises below evaluate this impact.

Table VIII conducts a two-part comparative statistics exercise on the calibrated model by shutting off uncertainty with respect to parts of the model's dynamics. The first panel, labeled Model 1-A, is for a model that shuts off uncertainty regarding only the jump shocks in the model, leaving on uncertainty regarding the diffusive parts of the dynamics. Thus, the jumps are not perturbed under the worst-case model. The second panel, Model 1-B, turns off all model uncertainty, so the investor has full confidence in the reference model. The table shows only the asset price data, since the cash flow dynamics being simulated are those of the reference model and therefore unaffected.

Table VIII
Comparative Statics Results

This table presents the results of the comparative statics exercise for the model of Table IV. The two panels alter the model in Table IV by successively turning off uncertainty towards aspects of the model. Model 1-A eliminates uncertainty with regards to the jump components of the model, but leaves uncertainty with respect to the diffusion dynamics. Model 1-B turns off all model uncertainty ($\varphi = 0$), so the agent has full confidence in the reference model. The reported percentiles are based on 10,000 simulations. Symbols are defined in Tables V and VI.

	Data				Model 1-A			Model 1-B		
	2007.6		2009.12		5%	50%	95%	5%	50%	95%
Returns										
$E[r_m - r_f]$	5.41	(2.09)	4.93	(2.08)	0.74	4.04	7.25	−2.74	0.76	4.11
$E[r_f]$	0.82	(0.35)	0.79	(0.34)	1.25	1.69	2.09	1.32	1.75	2.15
$\sigma(r_m - r_f)$	19.48	(2.35)	20.15	(2.24)	15.54	17.79	20.50	15.71	17.98	20.49
$\sigma(r_f)$	1.89	(0.17)	1.89	(0.16)	0.61	0.83	1.13	0.60	0.82	1.11
$skew(r_m - r_f)$ (M)	−0.43	(0.54)	−0.46	(0.52)	−0.41	−0.06	0.24	−0.24	0.03	0.33
$kurt(r_m - r_f)$ (M)	9.93	(1.26)	9.63	(1.23)	3.67	4.62	7.16	3.56	4.32	5.94
Variance Premium										
$E[VP]$	11.05	(1.07)	10.55	(2.62)	0.58	0.83	1.34	0.07	0.09	0.15
$\sigma(VP)$	7.63	(1.09)	8.47	(1.09)	0.35	0.78	1.68	0.04	0.09	0.19
$\sigma(\text{var}_t(r_m))$	17.41	(2.21)	31.00	(8.36)	8.26	13.55	25.93	8.19	12.96	24.84
$\sigma(\text{var}_t^Q(r_m))$	23.97	(3.15)	36.58	(8.07)	8.48	14.18	27.40	8.22	13.04	25.00
$\beta(1)$	0.74	(0.34)	0.80	(0.36)	−4.58	7.33	21.89	−97.03	25.20	153.24
$R^2(1)$	1.44	(1.48)	1.72	(1.64)	0.01	0.93	5.76	0.01	0.51	4.03
$\beta(3)$	0.87	(0.25)	0.99	(0.25)	−4.43	6.28	18.27	−89.46	19.79	129.58
$R^2(3)$	6.15	(4.50)	6.93	(4.53)	0.02	2.09	13.10	0.01	1.15	9.06

The top panel for Model 1-A shows that eliminating uncertainty with respect to the jump shocks reduces the median equity premium, though it remains significant. The other moments in the top panel are not greatly affected. Return volatility is reduced a little while the magnitudes of the monthly return skewness and kurtosis decrease somewhat more substantially. Note that for model 1-A the statistical "distance" between the worst-case and reference models is the same as before (i.e., the relative entropy of the worst-case model remains the same). For Model 1-B, where the relative entropy is zero and the worst-case model equals the reference model, the median equity premium is quite small.

The bottom panel shows the impact of model uncertainty on the variance premium. The results for Model 1-A show that eliminating uncertainty about jumps greatly reduces the size and volatility of the variance premium. Approximately 57% of the entropy of the worst-case model under the benchmark calibration is "allocated" to the jump perturbations. Without these channels, the average variance premium for Model 1-A is reduced by almost an order of magnitude and its simulation 95th percentile is nowhere near the data estimates. Without any model uncertainty, the average variance is reduced further until it is close to zero. Turning to the predictive regressions, the predictive R^2

of the variance premium is reduced successively in the two models. When the jump intensity is not directly amplified under the worst-case model, as under Model 1-A, the jumps' importance decreases along with the predictive power of the variance premium. For Model 1-B, where there is no model uncertainty, the influence of q_t^2 on the equity risk premium is very small and so is its predictive power. This is apparent in the diminished median R^2s shown in the table. Finally, as the variance premium becomes very small, the predictive regression coefficients become unstable. If the variance premium was this small empirically, it would likely be obscured by estimation noise.

The results from the table are augmented by Figure 3, and also the Internet Appendix, which plot the implied volatility curves for Model 1-B. The figures show that the model-generated implied volatility curves are very flat and that the model does not capture the steep skew in implied volatility. This is particularly apparent at the 1-month maturity, but is clearly also true for 3 and 12 months.

As an additional exercise, I shut off all model uncertainty (as for Model 1-B) and raise risk aversion to the point where the equity premium matches that of the benchmark calibration. The risk aversion required for this is 13.3. For such a configuration the mean variance premium across simulations has a median value of 4.23, and the median standard deviation of the variance premium is 4.00. These numbers are only about half of their benchmark counterparts. Thus, in this case a high risk aversion alone is insufficient for jointly capturing the variance and equity premiums. Moreover, a risk aversion of 13.3 already exceeds the value 10, often considered the upper end of the reasonable range for risk aversion, and is substantially higher than the risk aversion of 5 used in the benchmark calibration. Hence, for the benchmark calibration, model uncertainty, especially with regards to the jump shocks, is an important component for reconciling asset prices with fundamentals and a reasonable level of risk aversion.

Finally, Table IX looks at further comparative statics of the benchmark calibration, obtained by changing φ and hence the detection-error probability of the worst-case model. The table shows the results for five values of φ, in increasing order, with the middle column (III) setting φ equal to its benchmark calibration value. Recall that increasing φ raises the level of model uncertainty, increasing the distance between the worst-case and reference models and hence *decreasing* the probability of detection error. For each φ the top rows of the table show the detection-error probability of the worst-case model for data samples of length 240, 480, and 720 months. The bottom panels report asset pricing moments. Comparing across columns, the general patterns are clear. The equity premium increases and the risk-free rate decreases, as greater uncertainty makes the investor more concerned about lower cash flow growth rates and increased diffusive volatility and jump shocks. The joint impact of this uncertainty on variance expectations is shown by the variance premium, which increases sharply in the table. As the table shows, there is also a corresponding, but smaller, increase in the volatility of conditional variance. This results because, when the overall level of uncertainty is higher, variation in

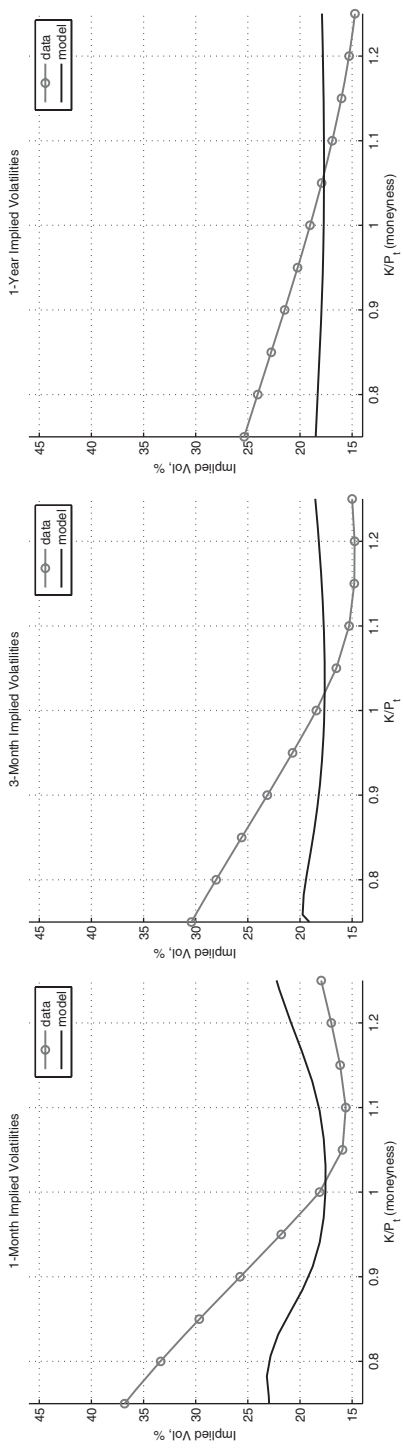

Figure 3. One-, 3-, and 12-month implied volatilities: No-uncertainty model and data. The figure plots comparisons of empirical and model-based implied volatilities for 1-, 3-, and 12-month maturities for Model 1-B, used for the comparative statics exercise in Table VIII. Strikes are expressed in moneyness (Strike Price/Spot Price). The empirical curves are means of daily implied volatilities for S&P 500 index options for the period 1999.10 to 2008.6. The model-based curves are calculated for option prices obtained when the model's state vector is set equal to its unconditional mean.

Table IX
Changing the Detection Error Probability

This table presents a comparison of the results for different levels of detection error probability for the model of Table IV. The detection error probability is changed by varying the parameter φ and is reported in the first three rows for samples of length $T = 240$, 480, and 720 months. The reported values for the moments are the medians from 10,000 simulations of the model. The values of φ corresponding to columns I to V are 0.0024, 0.0038, 0.0048, 0.0053, and 0.0058, respectively. Symbols are defined in Tables V and VI.

Statistic	I (50%)	II (50%)	III (50%)	IV (50%)	V (50%)
Sample Size / D.E. Prob.					
$T = 240$ months	0.323	0.282	0.261	0.252	0.246
$T = 480$ months	0.244	0.192	0.164	0.152	0.142
$T = 720$ months	0.194	0.137	0.109	0.097	0.088
$E[r_m - r_f]$	3.90	5.18	6.29	7.16	8.51
$E[r_f]$	1.69	1.62	1.52	1.45	1.33
$skew(r_m - r_f)$ (M)	−0.06	−0.16	−0.27	−0.36	−0.56
$kurt(r_m - r_f)$ (M)	4.62	5.08	5.73	6.21	7.50
$E[VP]$	1.85	4.29	8.17	11.92	19.37
$\sigma(VP)$	1.75	4.04	7.53	11.10	18.31
$\sigma(\text{var}_t(r_m))$	13.41	14.07	14.97	15.83	17.15
$\sigma(\text{var}_t^Q(r_m))$	14.77	17.46	21.69	25.99	34.42
$R^2(1)$	0.98	1.40	1.74	2.04	2.72
$R^2(3)$	2.23	3.08	3.91	4.63	6.27

the drivers of uncertainty also induce a stronger reaction in equity prices. The volatility of uncertainty shocks is then a more important driver of conditional variance and increases its overall volatility. When the uncertainty level is subject to jump shocks, as is the case here, then this also leads to a fatter tail in returns. This is apparent in the table, where, for higher levels of uncertainty, there is a more negative return skewness and greater kurtosis. In addition, the greater impact of uncertainty variation on equity prices leads to greater predictability of equity returns by the variance premium, as evidenced in the last two lines of the table.

VII. Conclusion

This paper builds an equilibrium model where the representative investor is concerned about a rich set of alternative models of the economy. The investor's attempt to hedge his model uncertainty optimally focuses his concerns on aspects of his model for which an error would be harmful to his utility while also being difficult to detect in the data. I solve for the investor's worst-case model and for equilibrium asset prices. I show that model uncertainty concerns help to capture the large premium contained in the prices of equity index options, and that variation in this option premium reflects variation in the level of model

uncertainty. This implies that the option premium should predict excess stock returns, consistent with the empirical results I present.

I present a calibrated specification of this model and show that with a statistically reasonable level of model uncertainty it is able to jointly capture salient moments of cash flows, equity returns, the risk-free rate, variance dynamics, and the large equity index option premium. Moreover, the model generates a realistic implied volatility surface. I demonstrate that model uncertainty has a large impact on the model's asset prices, especially the variance premium, which is impacted strongly by uncertainty concerns about the frequency and magnitude of jump shocks.

Introspection suggests that Knightian uncertainty is an important factor in decision making. Hence, model uncertainty can be an important channel in macroeconomic and financial models. There has also been a lot of discussion suggesting that Knightian uncertainty, the "unknown unknowns," was an influential factor in the recent economic crises. The plots of the model uncertainty proxy and estimated option premium presented in this paper are consistent with this notion and closely track the timeline of the crises. It therefore seems fruitful for future research to expand on the work here and in related studies to improve our understanding of the role of model uncertainty in macrofinance models.

Initial submission: April 27, 2010; Final version received: April 25, 2013
Editor: Campbell Harvey

Appendix A: Derivation of Relative Entropy Growth

As discussed in the main text, the relative entropy growth rate is defined as

$$R(\eta_t) = \frac{d}{ds}_{|s=0} E_t^{\eta}[\ln \eta_{t+s}].$$

Since $\eta_t = \eta_t^{dZ}\eta_t^{J}$, we can compute the relative entropy growth rate of the diffusion and jump parts separately and then add them. Furthermore, we can determine $R(\eta_t^{J})$ by adding the relative entropy growth rates of the component jumps.

To determine $R(\eta_t^{dZ})$, write $\ln \eta_t^{dZ}$ in terms of Z_t^{η} rather than Z_t (see Internet Appendix Section III. We then have that $\ln \eta_{t+s}^{dZ} = \frac{1}{2}\int_t^{t+s} h_u^T h_u du + \int_t^{t+s} h_u^T dZ_u^{\eta}$. Since the second term is a martingale under $P(\eta)$, its time-t expectation is zero. It is thus clear that

$$R(\eta_t^{dZ}) = \frac{1}{2} h_t^T h_t.$$

Now consider $R(\eta_t^{J_i})$, where η^{J_i} is the term in η_t that changes the probability law for the gamma-distributed jumps (from Internet Appendix Section III). To find $R(\eta_t^{J_i})$, recall two facts: (1) $l_t^{\eta} = \exp(a) l_t$ and (2) $E_t^{\eta}[\xi_i] = \frac{k\theta}{1-\theta b}$. Using these

facts, a straightforward calculation shows that

$$R\left(\eta_t^{J_i}\right) = \exp(a) l_t \left(a + b\frac{k\theta}{1-\theta b} + k\ln(1-\theta b) - 1\right) + l_t.$$

Finally, consider $R(\eta_t^{J_k})$, where η^{J_k} is the term in η_t that changes the probability law for the normally distributed jumps (from Internet Appendix Section III). Straightforward calculations and algebraic simplification show that

$$R\left(\eta_t^{J_k}\right) = \exp(a) l_t \left(a + \frac{1}{2}\frac{\Delta\mu^2}{\sigma^2} + \frac{1}{2}s_\sigma - \frac{1}{2}\ln s_\sigma - \frac{3}{2}\right) + l_t.$$

Appendix B: HJB Equation Derivation

Substituting (6) into the aggregator (5) gives

$$f(C_t, J_t) = \delta \exp(\gamma g(\tilde{Y}_t) + \gamma \ln C_t)\left[\frac{\exp(-\rho g(\tilde{Y}_t)) - 1}{\rho}\right].$$

Using Ito's lemma, we have that

$$E_t[dJ^c] = J_Y^T \mu(Y_t)dt + \frac{1}{2}\mathbf{tr}[J_{YY}\Sigma(Y_T)\Sigma(Y_t)^T]dt,$$

where the superscript "c" denotes the continuous (diffusive) part of the dynamics and **tr** denotes the trace operator. For notational convenience, let $G(Y_t) = g(\tilde{Y}_t) + \ln C_t$, so that we can write $J(Y_t) = \exp(\gamma G(Y_t))/\gamma$. Furthermore, let $G_Y(Y_t)$ and $G_{YY}(Y_t)$ denote the gradient and Hessian matrix of $G(Y_t)$. Then $J_Y = \exp(\gamma G(Y_t))G_Y(Y_t)$ and $J_{YY} = \exp(\gamma G(Y_t))\left[\gamma G_Y G_Y^T + G_{YY}\right]$. The Hamilton-Jacobi-Bellman (HJB) equation (4) can then be rewritten as

$$\begin{aligned} 0 = \min_{P(\eta_t)} & \frac{\delta}{\rho}\exp(\gamma G(Y_t))[\exp(-\rho g(\tilde{Y}_t)) - 1] + \exp(\gamma G(Y_t))G_Y^T \mu(\tilde{Y}_t) \\ & + \exp(\gamma G(Y_t))G_{Y_2}^T \Sigma_{2,t} h_{2,t} + \frac{1}{2}\exp(\gamma G(Y_t))\mathbf{tr}\left[\gamma G_Y^T \Sigma_t \Sigma_t^T G_Y + G_{YY}\Sigma_t\Sigma_t^T\right] \\ & + \frac{1}{\gamma}E_t^\eta[\exp(\gamma G(Y_{t_-} + \xi_t \cdot dN_t)) - \exp(\gamma G(Y_{t_-}))] \\ & \text{s.t.} \quad R(\eta_t) \le \varphi q_t^2, \end{aligned} \tag{B1}$$

where the arguments of the derivatives have been omitted to reduce clutter. Note that the term $\Sigma_{2,t}h_{2,t}$ is due to the drift perturbation under $P(\eta)$ and that G_{Y_2} just denotes the derivative of $G(Y)$ with respect to $Y_{2,t}$.

Appendix C: Proof of Proposition 1

The first step is to substitute the conjecture $g(\tilde{Y}_t) = A_0 + A'\tilde{Y}_t$ into the expanded HJB equation (B1) derived in Appendix B. To facilitate this, let $\hat{A}$ denote the vector obtained by augmenting A with a component equal to one, so that $\delta_c'\hat{A} = 1$. Then we can concisely write $G(Y_t) = A_0 + \hat{A}'Y_t$. Moreover, we see that $G_Y = \hat{A}$ and $G_{YY} = 0$. In addition, let $\hat{A} = [\hat{A}_1, \hat{A}_2]'$ be partitioned in the same way as $Y = [Y_1, Y_2]'$. Then we can write $G_{Y_2} = \hat{A}_2$. We can substitute in for the diffusion terms in (B1). For the jump terms, we have

$$\begin{aligned}
&\frac{1}{\gamma}E_t^{\eta}[\exp(\gamma G(Y_{t_-} + \xi_t \cdot dN_t)) - \exp(\gamma G(Y_{t_-}))] \\
&\quad = \frac{\exp(\gamma G(Y_t))}{\gamma}E_t^{\eta}[\exp(\gamma \hat{A}'(\xi_t \cdot dN_t)) - 1] \\
&\quad = \frac{\exp(\gamma G(Y_t))}{\gamma}l_t^{\eta'}(\psi^{\eta}(\gamma\hat{A}) - 1),
\end{aligned}$$

where, as before, $\psi^{\eta}(\gamma\hat{A})$ denotes the stacked vector of $P(\eta)$-measure moment-generating functions evaluated componentwise at the vector $\gamma\hat{A}$. I write out the functional form of these moment-generating functions for the two types of jumps considered, normal jumps and gamma jumps. Let ξ_i and ξ_k be the gamma and normally distributed jumps, respectively. As shown in Internet Appendix Section III under $P(\eta)$ we have $\xi_i \sim \Gamma(k, \frac{\theta}{1-\theta b})$ and $\xi_k \sim \mathcal{N}(\mu + \Delta\mu, \sigma^2 s_\sigma)$. Therefore, $\psi_i^{\eta}(\gamma\hat{A}_i) = (1 - \frac{\theta\gamma\hat{A}_i}{1-\theta b})^{-k}$ and $\psi_k^{\eta}(\gamma\hat{A}_k) = \exp(\gamma\hat{A}_k(\mu + \Delta\mu) + \frac{1}{2}\gamma^2\hat{A}_k^2\sigma^2 s_\sigma)$.

Substituting the derived expressions into (B1) and factoring out the term $\exp(\gamma G(Y_t))$ gives

$$\begin{aligned}
0 = \min_{P(\eta_t)} \exp(\gamma G(Y_t)) \times \Bigg\{ &\frac{\delta}{\rho}[\exp(-\rho g(\tilde{Y}_t)) - 1] + \hat{A}'\mu(\tilde{Y}_t) + \hat{A}_2{}'\Sigma_{2,t}h_{2,t} \\
&+ \frac{1}{2}\gamma\hat{A}'\Sigma_t\Sigma_t^T\hat{A} + \frac{1}{\gamma}l_t^{\eta'}(\psi^{\eta}(\gamma\hat{A}) - 1)\Bigg\} \\
\text{s.t.} \quad & R(\eta_t) \le \varphi q_t^2.
\end{aligned} \tag{C1}$$

We can now proceed with the minimization. Let λ_t be the Lagrange multiplier on the constraint $R(\eta_t) \le \varphi q_t^2$. The functional form of $R(\eta_t)$ in terms of the worst-case parameters is given by the sum of the expressions in Appendix A. The first-order conditions for the minimization are taken with respect to the the worst-case model parameters $(h_{2,t}, a, \Delta u, s_\sigma, b)$. They are:

$$\begin{aligned}
&\text{FOC}(h_{2,t}) : \exp(\gamma G(Y_t))\Sigma_{2,t}'\hat{A}_2 + \lambda_t h_{2,t} = 0 \\
&\text{FOC}(\Delta\mu) : \exp(\gamma G(Y_t))\frac{1}{\gamma}l_{t,k}^{\eta}\frac{\partial}{\partial\Delta\mu}\psi^{\eta}(\gamma A_k) + \lambda_t\frac{\partial}{\partial\Delta\mu}R(\eta_t^{J_k}) = 0 \\
&\text{FOC}(s_\sigma) : \exp(\gamma G(Y_t))\frac{1}{\gamma}l_{t,k}^{\eta}\frac{\partial}{\partial s_\sigma}\psi^{\eta}(\gamma A_k) + \lambda_t\frac{\partial}{\partial s_\sigma}R(\eta_t^{J_k}) = 0
\end{aligned}$$

$$\text{FOC}(b): \exp(\gamma G(Y_t))\frac{1}{\gamma} l_{t,i}^{\eta}\frac{\partial}{\partial b}\psi^{\eta}(\gamma A_i) + \lambda_t \frac{\partial}{\partial b} R(\eta_t^{J_i}) = 0$$

$$\text{FOC}(a): \exp(\gamma G(Y_t))\frac{1}{\gamma}\ \exp(a) l_t{}'(\psi^{\eta}(\gamma \hat{A}) - 1) + \lambda_t \frac{\partial}{\partial a} R(\eta_t) = 0.$$

Moreover, the entropy constraint binds, which adds the equation $R(\eta_t) = \varphi q_t^2$. For a given value of the vector A, these equations fix the values of the worst-case parameters. To solve the system, I conjecture that $\lambda_t = \tilde{\lambda}\exp(\gamma G(Y_t))$, where $\tilde{\lambda} > 0$ is a constant that defines the part of the Lagrange multiplier that is state-*independent*. By substituting this conjecture into the above system, I can divide through by $\exp(\gamma G(Y_t))$ in each of the equations and cancel the state-dependent terms $\exp(\gamma G(Y_t))$ and λ_t, leaving the constant $\tilde{\lambda}$.

Solving the first equation, we obtain $h_{2,t} = -\frac{1}{\tilde{\lambda}}\Sigma_{2,t}'\hat{A}_2$. This implies that the perturbation in the model drift is $\Sigma_{2,t}h_{2,t} = -\frac{1}{\tilde{\lambda}}H_q A_2 q_t^2$ and the contribution of h_t to $R(\eta_t)$ is $\frac{1}{2}h_t'h_t = \frac{1}{2}\frac{1}{\tilde{\lambda}^2}A_2' H_q A_2 q_t^2$. Thus, the drift perturbation is proportional to q_t^2 and the perturbation's cost in terms of entropy is also proportional to q_t^2. This expression for the drift's entropy cost can now be substituted into the entropy constraint and we can eliminate the first-order condition for $h_{2,t}$ from the above system.

The remaining issue for solving this system is the dependence of a number of terms on the level of q_t^2. As just shown, the term $\frac{1}{2}h_t'h_t$ in the entropy constraint is proportional to q_t^2. Furthermore, the jump intensity l_t, which is proportional to q_t^2, appears in the other first-order conditions. It appears both explicitly on the left side of each equation and implicitly in the partial derivatives of R (the expression for R is in Appendix A). This means that all of the remaining equations are in fact proportional to q_t^2 and so q_t^2 can be canceled out from all of them. The resulting equation system is now completely state-*independent* and has as its solution the *constant* vector of worst-case parameters $(a, \Delta u, s_\sigma, b)$ and $\tilde{\lambda}$.

The remaining step is to find the equations determining the value of the vector A and in the process verify the conjectured solution for $g(\tilde{Y}_t)$. Toward that end, I approximate the term $\exp(-\rho g(\tilde{Y}_t))$ in (C1) via a log-linearization. As the Internet Appendix shows, $\exp(-\rho g(\tilde{Y}_t))$ is just $\frac{1}{\delta}$ times the equilibrium consumption-wealth ratio. I follow Campbell et al. (2004) and log-linearize it around the unconditional mean of the equilibrium log consumption-wealth ratio as follows: $\exp(-\rho g(\tilde{Y}_t)) \approx \kappa_0 + \kappa_1 \rho g(\tilde{Y}_t)$, where $\kappa_1 = -\exp(-\rho E[g(\tilde{Y}_t)])$ and $\kappa_0 = -\kappa_1(1 + \rho E[g(\tilde{Y}_t)]) = -\kappa_1(1 - \ln(-\kappa_1))$. Note that the values of κ_0 and κ_1 are endogenous to the equilibrium solution of the model. As Campbell et al. (2004) point out, this property of the approximation is important in obtaining accuracy over a range of values for ψ and γ. Furthermore, L'Hospital's rule gives that $\lim_{\rho\to 0}\frac{\exp(-\rho g(\tilde{Y}_t))-1}{\rho} = -g(\tilde{Y}_t) = \lim_{\rho\to 0}\frac{\kappa_0+\kappa_1\rho g(\tilde{Y}_t)-1}{\rho}$, so the approximation to the PDE becomes exact as $\rho \to 0$ ($\psi \to 1$). Substituting in this approximation and the functional form of $g(\tilde{Y}_t)$, the first term in (C1) becomes $\delta(\frac{\kappa_0-1}{\rho} + \kappa_1 A_0) + \delta\kappa_1 A'\tilde{Y}_t$.

Given the worst-case model parameters, (C1) must hold for all values of $\tilde{Y}_t$. The solution to this equation determines the values of A and A_0 and verifies the conjectured solution for $g(\tilde{Y}_t)$. As $\tilde{Y}_t$ is $(n-1)$-dimensional, the system of equations is n-dimensional, one equation for each of the $(n-1)$ elements in $\tilde{Y}_t$ and one for the constant terms. Expanding out (C1) (and using the log-linearization), we get the following system:

$$
\begin{aligned}
0 &= \frac{\delta}{\rho}(\kappa_0 + \kappa_1 \rho A_0 - 1) + \hat{A}^T \mu \\
0 &= \tilde{Y}_t^T A \delta \kappa_1 + \tilde{Y}_t^T \tilde{\mathcal{K}}^T \hat{A} - \frac{1}{\tilde{\lambda}}(\hat{A}_2^T H_q \hat{A}_2) q_t^2 + \frac{1}{2}\gamma \hat{A}^T \Sigma_t \Sigma_t^T \hat{A} \\
&\quad + \frac{1}{\gamma} l_1^{\eta'}(\psi^{\eta}(\gamma \hat{A}) - 1) q_t^2.
\end{aligned} \tag{C2}
$$

Combining these equations with the first-order conditions above and the equations defining κ_0 and κ_1 gives the system of equations that defines the equilibrium solution. The solution can be found numerically and verifies the conjectured functional form for the value function.

REFERENCES

Andersen, Torben G., Tim Bollerslev, and Francis X. Diebold, 2007, Roughing it 4p: Including jump components in the measurement, modeling, and forecasting of return volatility, *Review of Economics and Statistics* 89, 701–720.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Heiko Ebens, 2001, The distribution of stock return volatility, *Journal of Financial Economics* 61, 43–76.

Anderson, Evan W., Eric Ghysels, and Jennifer L. Juergens, 2009, The impact of risk and uncertainty on expected returns, *Journal of Financial Economics* 94, 233–363.

Anderson, Evan W., Lars P. Hansen, and Thomas J. Sargent, 2003, A quartet of semigroups for model specification, robustness, prices of risk, and model detection, *Journal of the European Economic Association* 1, 68–123.

Avramov, Doron, 2002, Stock return predictability and model uncertainty, *Journal of Financial Economics* 64, 423–458.

Avramov, Doron, 2004 Stock return predictability and asset pricing models, *Review of Financial Studies* 17, 699–738.

Backus, David, Mikhail Chernov, and Ian Martin, 2011, Disasters implied by equity index options, *Journal of Finance* 66, 1969–2012.

Bakshi, Gurdip, and Nikunj Kapadia, 2003, Delta-hedged gains and the negative volatility risk premium, *Review of Financial Studies* 16, 527–566.

Bansal, Ravi, Robert F. Dittmar, and Christian Lundblad, 2005, Consumption, dividends, and the cross-section of equity returns, *Journal of Finance* 60, 1639–1672.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2007, Risks for the long run: Estimation and inference, Working paper, The Wharton School, University of Pennsylvania.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An empirical evaluation of the long-run risks model for asset prices, *Critical Finance Review* 1, 183–221.

Bansal, Ravi, and Amir Yaron, 2004, Risks for the long run: A potential resolution of asset pricing puzzles, *Journal of Finance* 59, 1481–1509.

Barillas, Francisco, Lars Hansen, and Thomas Sargent, 2009, Doubts or variability? *Journal of Economic Theory* 144, 2388–2418.

Barro, Robert J., 2006, Rare disasters and asset markets in the twentieth century, *Quarterly Journal of Economics* 121, 823–867.

Bates, David S., 2003, Empirical option pricing: A retrospection, *Journal of Econometrics* 116, 387–404.

Beeler, Jason, and John Y. Campbell, 2012, The long-run risks model and aggregate asset prices: An empirical assessment, *Critical Finance Review* 1, 141–182.

Ben-David, Itzhak, John R. Graham, and Campbell R. Harvey, 2010, Managerial miscalibration, Working paper, The Ohio State University and Duke University.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2011, Explaining asset pricing puzzles associated with the 1987 market crash, *Journal of Financial Economics* 101, 552–573.

Bollerslev, Tim, Michael Gibson, and Hao Zhou, 2011, Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities, *Journal of Econometrics* 160, 102–118.

Bollerslev, Tim, George Tauchen, and Hao Zhou, 2009, Expected stock returns and variance risk premia, *Review of Financial Studies* 22, 4463–4492.

Brennan, Michael J., and Yihong Xia, 2001, Stock price volatility and equity premium, *Journal of Monetary Economics* 47, 249–283.

Brevik, Frode, 2008, State uncertainty aversion and the term structure of interest rates, Working paper, VU Amsterdam.

Britten-Jones, Mark, and Anthony Neuberger, 2000, Option prices, implied price processes, and stochastic volatility, *Journal of Finance* 55, 839–866.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model specification and risk premiums: Evidence from futures options, *Journal of Finance* 62, 1453–1490.

Campbell, John Y., George Chacko, Jorge Rodriguez, and Luis M. Viceira, 2004, Strategic asset allocation in a continuous-time VAR model, *Journal of Economic Dynamics and Control* 28, 2195–2214.

Campbell, John Y., and John H. Cochrane, 1999, By force of habit: A consumption-based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205–255.

Campbell, John Y., and Robert Shiller, 1988, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661–676.

Carr, Peter, and Dilip Madan, 1999, Option valuation using the fast Fourier transform, *Journal of Computation Finance* 2, 61–73.

Carr, Peter, and Liuren Wu, 2009, Variance risk premiums, *Review of Financial Studies* 22, 1311–1341.

Coval, Joshua D., and Tyler Shumway, 2001, Expected option returns, *Journal of Finance* 56, 983–1010.

Demeterfi, Kresimir, Emanuel Derman, Michael Kamal, and Joseph Zou, 1999, A guide to volatility and variance swaps, *Journal of Derivatives* 6, 9–32.

Drechsler, Itamar, and Amir Yaron, 2011, What's vol got to do with it, *Review of Financial Studies* 24, 1–45.

Du, Du, 2010, General equilibrium pricing of options with habit formation and event risks, *Journal of Financial Economics* 99, 400–426.

Duffie, Darrell, and Larry G. Epstein, 1992, Stochastic differential utility, *Econometrica* 60, 353–394.

Duffie, Darrell, Jun Pan, and Kenneth J. Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Epstein, Larry G., and Martin Schneider, 2003, Recursive multiple-priors, *Journal of Economic Theory* 113, 1–31.

Epstein, Larry G., and Stanley E. Zin, 1989, Substitution, risk aversion, and the intertemporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

Eraker, Bjorn, 2004, Do equity prices and volatility jump? Reconciling evidence from spot and option prices, *Journal of Finance* 56, 1367–1403.

Eraker, Bjorn, and Ivan Shaliastovich, 2008, An equilibrium guide to designing affine pricing models, *Mathematical Finance* 18, 519–543.

Foresi, Silverio, and Liuren Wu, 2005, Crash-o-phobia: A domestic fear or worldwide concern?, *Journal of Derivatives* 13, 8–21.

Hansen, Lars, and Thomas Sargent, 2010, Fragile beliefs and the price of model uncertainty, *Quantitative Economics* 1, 129–162.

Hansen, Lars P., and Thomas J. Sargent, 2008, *Robustness* (Princeton University Press, Princeton, NJ).

Hansen, Lars P., Thomas J. Sargent, Gauhar A. Turmuhambetova, and Noah Williams, 2006, Robust control and model misspecification, *Journal of Economic Theory* 128, 45–90.

Hansen, Peter R., and Asger Lunde, 2006, Realized variance and market microstructure noise, *Journal of Business and Economics Statistics* 24, 127–161.

Jiang, George, and Yisong Tian, 2005, Model-free implied volatility and its information content, *Review of Financial Studies* 18, 1305–1342.

Kleshchelski, Isaac, and Nicolas Vincent, 2007, Robust equilibrium yield curves, Working paper, Northwestern.

Liu, Jun, Jun Pan, and Tan Wang, 2005, An equilibrium model of rare-event premia and its implication for option smirks, *Review of Financial Studies* 18, 131–164.

Maenhout, Pascal J., 2004, Robust portfolio rules and asset pricing, *Review of Financial Studies* 17, 951–983.

Maenhout, Pascal J., 2006, Robust portfolio rules and detection error probabilities for a mean-reverting risk premium, *Journal of Economic Theory* 128, 136–163.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding predictability, *Journal of Political Economy* 112, 1–47.

Pan, Jun, 2002, The jump-risk premia implicit in options: Evidence from an integrated time- series study, *Journal of Financial Economics* 63, 3–50.

Pástor, Ľuboš, 2000, Portfolio selection and asset pricing models, *Journal of Finance* 55, 179–223.

Pástor, Ľuboš, and Robert Stambaugh, 2000, Comparing asset pricing models: An investment perspective, *Journal of Financial Economics* 56, 335–381.

Pástor, Ľuboš, and Robert Stambaugh, 2011, Are stocks really less volatile in the long run?, *Journal of Finance* 67, 431–478.

Rubinstein, Mark, 1994, Implied binomial trees, *Journal of Finance* 69, 771–818.

Shaliastovich, Ivan, 2011, Learning, confidence, and option prices, Working paper, Wharton School, University of Pennsylvania.

Singleton, Kenneth, 2006, *Empirical Dynamic Asset Pricing* (Princeton University Press, Princeton, NJ).

Todorov, Viktor, 2010, Variance risk premium dynamics: The role of jumps, *Review of Financial Studies* 23, 345–383.

Trojani, Fabio, and Alessandro Sbuelz, 2008, Asset prices with locally-constrained-entropy recursive multiple priors utility, *Journal of Economic Dynamics and Control* 32, 3695–3717.

Ulrich, Maxim, 2013, Inflation ambiguity and the term structure of U.S. Government bonds, *Journal of Monetary Economics* 60, 295–309.

Uppal, Raman, and Tan Wang, 2003, Model misspecification and underdiversification, *Journal of Finance* 58, 2465–2486.

Veronesi, Pietro, 2000, How does information quality affect stock returns? *Journal of Finance* 55, 807–837.