Uncertainty, Time-Varying Fear, and Asset Prices *

Itamar Drechsler†

March 2009

Abstract

I argue that time-varying Knightian uncertainty regarding economic fundamentals plays a central role in accounting for the equity premium, return volatility and the large, volatile variance premium embedded in equity index option prices. I build a general equilibrium framework that incorporates time-varying Knightian uncertainty about diffusive and jump shocks to the level and volatility of long-run cash-flow growth rates. A calibrated model is shown to capture the variance premium and option skew while simultaneously matching the moments of cash-flows and stock returns. The model indicates that fluctuations in the variance premium strongly reflect changes in the level of Knightian uncertainty and should predict monthly stock returns, consistent with recent empirical evidence.

*I thank my committee, Amir Yaron (Chair), Rob Stambaugh, and Stavros Panageas. I also thank Andy Abel, Joao Gomes, Philipp Illeditsch, Jakub Jurek, Richard Kihlstrom, Fei fei 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, and the 2009 WFA (San Diego) for helpful comments. I thank Nim Drexler and contacts at Citigroup and CSFB for over-the-counter options data.

†The Wharton School, University of Pennsylvania, idrexler@wharton.upenn.edu.

1 Introduction

When making decisions an investor faces both risk and uncertainty. Risk represents randomness that the investor understands and can model, while Knightian or model uncertainty represents ambiguity about the data-generating process itself.[1] This paper argues that the prices of index options are sensitive to investors' level of uncertainty and that uncertainty and its time-variation are central in allowing a general equilibrium model to jointly capture option prices, the equity premium, and the properties of cash-flows.

The prices of index options and the implied-volatility skew pose a considerable challenge to equilibrium asset pricing models. Studies have found that index options are priced with a considerable premium.[2] A measure of this is the variance premium, which is defined as the difference between the option-implied (VIX) and statistical (true) expectation of one-month return variance. By measuring the difference between risk-neutral and true expectations of variance, the variance premium captures the impact that pricing concerns have on the value of options. Empirically, the variance premium is found to be significant and systematically positive, implying that buyers of index options pay a large hedging premium. Its variation also appears to be related to expected equity returns, as recent studies have found that it is able to predict stock returns at monthly horizons.

Time-varying model uncertainty presents an intuitive potential explanation for the large magnitude and variation of the variance premium and the high option skew. Options provide investors with a natural protection against worrisome model specification concerns. An example of such concerns is underestimating the frequency or magnitude of jump shocks to important economic state variables, such as the long-run growth rates of cash-flows. As a result, time-variation in uncertainty concerns should be strongly reflected in option premia, which further implies that options provide a hedge to variation in the level of uncertainty itself.

In order to investigate the ability of uncertainty to quantitatively account for the variance premium and option prices, I build a framework that incorporates time-varying uncertainty about economic fundamentals into a representative agent equilibrium model that includes

[1] From now on, I shall take the terms uncertainty, ambiguity, and Knightian/model uncertainty to mean the same thing and use them interchangeably.

[2] See for example, Coval and Shumway (2001), Pan (2002), Bakshi and Kapadia (2003), Eraker (2004), and Carr and Wu (2007). Singleton (2006) contains a review of option-pricing studies.

both diffusive and jump risks, and I solve for asset and option prices. The representative agent in this economy has in mind a benchmark or reference model of the economy's dynamics that represents his best estimate of the data generating process. The agent is concerned that his reference model is misspecified and that the true model is actually in a set of alternative models that are statistically 'close' to the reference model. 'Close' means these alternative models are difficult to distinguish statistically based on historical data, so the agent's concerns about the reference model are reasonable. The level of uncertainty determines the size of the alternative set of models he worries about at a given time; when uncertainty increases the set of alternative models expands.[3]

Within this framework I conduct a quantitative calibration of an economy where the agent has uncertainty about the dynamics of long-run cash flow growth rates and the frequency and magnitude of related jump shocks. The calibration demonstrates that the model is able to capture the variance premium and option skew while simultaneously matching the equity premium and return volatility, riskfree-rate, and moments of consumption and dividends. The model also endogenously generates return predictability by the variance premium that is consistent with the data. It further predicts that option-implied quantities, such as the variance premium and VIX, should be better predictors of excess stock returns than statistical measures of variance, which is consistent with empirical findings.

A central challenge for general equilibrium asset pricing models that address these data is that cross-asset relationships make it difficult to account for the properties of the variance premium without implying an unrealistically high equity premium and return volatility or excessively volatile cash-flow processes. By amplifying fears of jump shocks to the level and volatility of long-run risks, uncertainty increases the prices of options. In addition, since time-varying uncertainty acts like a risk-factor, options earn a premium for hedging changes in uncertainty. This allows the model to match the data with a reasonable level of uncertainty and a relative risk aversion of 5.

The success of the model in reconciling a myriad set of cash-flow and asset pricing moments requires a flexible framework, which the paper builds and solves. The model that the agent considers includes a persistent growth rate (long-run risk), moderate jump shocks,

[3]This framework corresponds to the literature on Robust Control, which has been pioneered by Hansen and Sargent. See e.g. Anderson, Hansen, and Sargent (2003), Hansen, Sargent, Turmuhambetova, and Williams (2006), Hansen and Sargent (2007), and Hansen and Sargent (2008). The preferences are in the class of Recursive Multiple-Priors Utility of Epstein and Schneider (2003).

and stochastic volatility. In addition to allowing the model to be realistic enough for the calibration to match a large set of moments, this flexibility also allows model uncertainty to operate through multiple channels. The equilibrium solution endogenously determines the extent of specification concerns about different parts of the economic dynamics. The amount of concern is determined through a tradeoff between the damage a model specification error causes to lifetime utility and the difficulty of detecting it. The most important specification errors have large effects on utility but are difficult to detect. In the calibration, infrequent jump shocks related to long-run growth rates present prominent specification concerns. Such shocks have large cumulative effects on utility and are also difficult to detect, so underestimation of their frequency or size is a key specification concern.

Related Literature

This paper is related to a number of papers that study the variance premium and option prices. Bollerslev, Gibson, and Zhou (2008) and Bollerslev and Zhou (2007) also measure the variance premium using the difference between option-implied and realized variance measures of volatility. Both papers find that their measures have significant predictive power for stock returns at short horizons (a few months). Santa-Clara and Yan (2008) extract a measure of jump intensity from their option-pricing model and find it predicts stock returns. The paper is also related to option-pricing studies that confront their models with both physical and risk-neutral (i.e. price) data and conclude that jumps are necessary (e.g. Pan (2002), Eraker (2004), Broadie, Chernov, and Johannes (2007)). The model here shares that view. A big difference is that the model here is preference-based and derives prices starting from macroeconomic fundamentals. Benzoni, Collin-Dufresne, and Goldstein (2005) combine recursive preferences with a very large and rare shock to a persistent component in cash-flow growth rates to generate a steep implied-volatility smirk. Their model does not incorporate stochastic volatility and they do not analyze the variance premium. Eraker and Shaliastovich (2008) describe benefits of an equilibrium approach to option-pricing and build a pricing framework based on recursive preferences and affine dynamics that expands on Bansal and Yaron (2004) and Tauchen (2005) by including jump shocks and pricing options. Eraker (2008) estimates such a model, which captures the variance premium with very rare but large jumps in consumption volatility. In his model there is no persistent component in cash-flow growth rates, an important risk channel in both Bansal and Yaron (2004) and this paper. A very different paper that is also related is Anderson, Ghysels, and Juergens (2007),

which constructs measures of Knightian uncertainty using survey forecasts and finds that these measures are able to to predict stock returns in the time-series and the cross-section.

In its application of model uncertainty to explaining option prices, this paper is related to Liu, Pan, and Wang (2005) (LPW), who use uncertainty towards rare events to explain the smirk pattern in index options. There are, however, a number of significant differences with LPW. The environment in LPW is i.i.d, so it cannot address the conditional moments considered here, such as the volatility and return predictability of the variance premium, or the 'excess volatility' of returns. Second, the calibration in LPW is limited to only the equity premium and slope of the option smirk and does not consider other moments of equity returns, the risk-free rate, or properties of cash flows. Third, LPW model robustness towards 'rare-disasters', i.e. large, rare jumps in the aggregate endowment, while the calibration here focuses on jumps that occur (on average) every year or two and are small to moderate. Moreover, jump shocks enter the endowment only through a small, persistent component in growth rates and therefore do not cause immediate, large drops in aggregate consumption. Finally, the framework developed here allows for multiple state variables, uncertainty regarding both diffusions and jumps, and recursive utility.[4] The framework in this paper is also related to the work in Trojani and Sbuelz (2008), who specify a time-varying set of alternative models that depends on a state variable, and to Maenhout (2004), who solves for the equity premium in an economy with a robust agent that has recursive utility.

Finally, this paper is related to Drechsler and Yaron (2008), who build an extended long-run risks model with jump shocks that captures the size and predictive power of the variance premium. They demonstrate that the variance premium effectively reveals variation in the intensity of jump shocks, which accounts for its predictive power. This paper differs in its focus on Knightian uncertainty as a key component of the model. While time-variation in the risk of jump shocks is still the main driver of the variance premium, it arises from the combination of jump risks in the reference model and model uncertainty. Model uncertainty amplifies concerns about influential jump shocks, so that less is needed in terms of physical jumps. This feature enables the model in this paper to capture the equity premium and option prices with a relatively low risk aversion of 5. Finally, the model in this paper generates stochastic return volatility through two channels–stochastic cash-flow volatility

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

and stochastic uncertainty. This makes return volatility, the VIX, and variance premium be imperfectly correlated and allows the model to capture why the latter two quantities are superior predictors of equity returns.

2 Definitions and Data

The definitions of key terms is similar to those in Bollerslev and Zhou (2007) and follows related literature. I define the variance premium as the difference between the risk neutral and physical expectations of the market's total return variation. I focus on a one month variance premium, so the expectations are of total return variation between the current time, t, and one month forward, $t+1$. Thus, $vp_{t,t+1}$, the (one-month) variance premium at time t, is defined as $E_t^Q\left[\int_t^{t+1}\left(d\ln R_{m,s}\right)^2\right]$ - $E_t\left[\int_t^{t+1}\left(d\ln R_{m,s}\right)^2\right]$ where Q denotes the risk-neutral measure and $\ln R_{m,s}$ is the (log) return on the market.

Demeterfi, Derman, Kamal, and Zou (1999) and Britten-Jones and Neuberger (2000) show that, in the case that the underlying asset price is continuous, the risk neutral expectation of total return variance is equal to the value of a portfolio of European calls on the asset. Jiang and Tian (2005) and Carr and Wu (2007) show this result extends to the case where the asset is a general jump-diffusion. This approach to calculating risk-neutral variance expectations is model-free since the calculations do not depend on any particular model of options prices. The VIX Index is calculated by the Chicago Board Options Exchange (CBOE) using this model-free approach to obtain the risk-neutral expectation of total variation 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. 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. Below I refer to the resulting series as squared VIX.

As the definition of $vp_{t,t+1}$ indicates, in order to measure it one also needs conditional forecasts of total return variation under the physical measure. To obtain these forecasts I measure the total realized variation of the market, or realized variance, for the months in my sample. For a given month, this measure is created by summing the squared five-minute log returns on the S&P 500 futures over the whole month. I obtain the high frequency futures data used in the construction of the realized variance measure from TICKDATA. To get the conditional forecasts, I project the realized variance measure on the value of the squared VIX

at the end of the previous month and on lagged realized variance and use the resulting series of forecasts for realized variance. The forecast series serves as the proxy for the conditional expectation of total return variance under the physical measure. The difference between the risk neutral expectation, measured using the squared VIX, and the contemporaneous conditional forecast from the projection, gives the series of one-month variance premium estimates. The projection specification I use is the same as in Drechsler and Yaron (2008). See that paper for further details.

The data series for the VIX and realized variance measures covers the period January 1990 to March 2007. The main limitation on the length of the sample comes from the VIX, which is only published by the CBOE beginning in January of 1990. I also present a comparison of the empirical and model-based implied-volatility surfaces for S&P 500 index options. Daily data on the volatility surface was obtained from Citigroup and covers October 1999 to June 2008. The model calibration I conduct also requires data on consumption and dividends and I use the longest sample available (1930:2006). Per-capita consumption of non-durables 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, the measure of futures realized variance, and the variance premium measure (VP). Note that the mean of the variance premium is sizeable in comparison with that of the squared VIX and realized variance. It is also quite volatile. Since any difference between risk-neutral and physical expectations is due to pricing considerations, the large, positive variance premium indicates that the VIX and option prices embed a substantial pricing premia. Capturing this premia and its volatility is a major challenge for an equilibrium model. Note further that all three variance-related series display significant deviation from normality. The mean to median ratio is large, the skewness is positive and greater than 0, and the kurtosis is clearly much larger than 3.

Table II provides return predictability regressions. There are two sets of columns with regression estimates. The first set shows OLS estimates and the second set provides estimates from robust regressions. Robust regression performs estimation using an iterative reweighted least squares algorithm that downweights the influence of outliers on estimates but is nearly as statistically efficient as OLS in the absence of outliers. It provides a check that the results are not driven by outliers. 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, which is commonly used variable for predicting returns. As a univariate regressor, the variance premium can account for about 1.5-4.0% of the monthly return variation. The multivariate regressions lead to a substantial further increase in the R^2 – a feature highlighted in Bollerslev and Zhou (2007). In conjunction with the price-earnings ratio, the in-sample R^2 increases to over 13%. Note that in all cases the variance premium enters with a significant positive coefficient. This sign and magnitude will be shown to be consistent with the theory in this paper. Finally, note that the robust regression estimates agree both in magnitude and sign with the OLS estimates and in fact, some of the R-squares are even larger than their OLS counterparts.

The conditional forecasts of total return variation used in constructing $vp_{t,t+1}$ are based on a regression estimated using the whole data sample. The use of the full sample raises the possibility that these forecasts may somehow be materially effected by information not available at the time of the forecast. An alternative approach that avoids this is to estimate the variance forecasting regression on a rolling basis using only past data. Table III repeats the return predictability regressions using the measure of $vp_{t,t+1}$ obtained by using the variance forecasts from these rolling regressions. The first 24 months are used to initialize the rolling regression estimates, so the return forecasts begin in January 1992. A comparison with Table II shows that the results are largely unchanged and are in fact a somewhat stronger.

3 Model

The setting is an infinite-horizon, continuous-time exchange economy with a representative agent who has utility over consumption streams. This agent has in mind a benchmark or *reference* model of the economy that represents his best estimate of the economy's dynamics. However, the agent does not fully trust that his model is correct. His model uncertainty concerns cause him to worry that the true model lies in a set of *alternative* models that are difficult for him to reject based on the data. This set of models is 'close' to the reference model in the sense that they are statistically difficult to distinguish from the reference model.

The agent guards against model uncertainty by acting cautiously and evaluating his future prospects under the worst-case model in the alternative set of models. The details follow.

3.1 The Reference Model

This section describes the agent's reference model. Let Y_t denote the n-dimensional vector of state variables. The reference model dynamics follow a continuous-time affine jump-diffusion:[5]

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

where Z_t is an n-dimensional Brownian motion, $\mu(Y_t)$ is an n-dimensional vector, $\Sigma(Y_t)$ is $n \times n$-dimensional matrix and both ξ_t and N_t are n-dimensional vectors. The term $\xi_t \cdot dN_t$ denotes component-wise multiplication of the jump sizes in the random vector ξ_t and the vector of increments in the Poisson (counting) processes N_t. The Poisson arrivals are conditionally independent and arrive with a time-varying intensity given by the n-dimensional vector l_t. The jump sizes in ξ_t are assumed to be i.i.d through time and in the cross-section. To handle the jumps, it is convenient to specify their moment generating function (mgf). Let $\psi_k(u) = E[\exp(u\xi_k)]$ be the mgf of the random jump size ξ_k. It is convenient to stack the mgf's into a vector function denoted $\psi(u)$. Thus, for u, an n-dimensional vector, $\psi(u)$ is the vector with k-th component $\psi_k(u_k)$. It is also assumed that log consumption and dividend growth are linear in Y_t:

$$\begin{aligned} d\ln C_t &= \delta_c' dY_t \\ d\ln D_t &= \delta_d' dY_t \end{aligned}$$

For convenience $\ln C_t$ and $\ln D_t$ are included in Y_t, so δ_c and δ_d are just selection vectors.

It is assumed that the drift, diffusion and jump intensity functions have an affine structure. Specifically,

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

where μ and $\mathcal{K}$ are n and $n \times n$-dimensional respectively. It is further assumed that the

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

economy's dynamics under the reference model are independent of the level of consumption C_t. This standard asset pricing assumption leads to an equilibrium that is homogenous in the level of consumption. The assumption can be formalized as $\mathcal{K}\delta_c = 0$, i.e. the column corresponding to C_t is just 0. To simplify some of the later exposition, let $\tilde{\mathcal{K}}$ denote the $n \times (n-1)$ dimensional sub-matrix of $\mathcal{K}$ which excludes this column and let $\tilde{Y}_t$ be the sub-vector of Y_t that excludes C_t. The assumption can then be rewritten as $\mathcal{K}Y_t = \tilde{\mathcal{K}}\tilde{Y}_t$.

Let $Y = (Y_{1,t}, Y_{2,t})$ be a partition of the state vector. The diffusion covariance matrix has a block-diagonal form:

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

where the upper block corresponds to $Y_{1,t}$ and the lower block to $Y_{2,t}$. $\Sigma_{1,t}\Sigma_{1,t}'$ has a general affine form:

$$\Sigma_{1,t}\Sigma_{1,t}' = h + \sum_i H_i Y_{t,i}$$

Let q_t denote a particular state variables in Y_t. This variable will appear repeatedly throughout the model and has the important role of governing variation in the agent's level of uncertainty, as discussed below. 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 an n-dimensional vector.[6]

The partition of Y_t relates to which subset of the model dynamics the agent is uncertain about. I make the agent uncertain about only the dynamics of $Y_{2,t}$, which contains the variables the agent feels are difficult to detect or estimate.[7] Thus, the specification above makes q_t drive the volatility of shocks about which there is uncertainty. There are several motivations for this specification. First, it seems reasonable that the level of uncertainty should rise when there is an elevated risk of large shocks to important state variables, i.e. it seems plausible that the level of uncertainty be related to economic risk. The calibration

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

[7]Of course we can have $Y_{2,t} = Y_t$

section discusses some evidence on an empirical measure of uncertainty that is consistent with this idea. Finally, this specification also facilitates analytical tractability. To isolate the pure effects of variation in uncertainty in a stark manner, I also include a discussion at the end of the paper of a modified model where q_t is orthogonal to the other processes.

3.2 Alternative Models

This section describes the agent's set of alternative models. An alternative model is defined by its probability measure. A requirement of this probability measure is that it put positive probability on the same events as the reference model's probability measure (i.e. they are equivalent measures). Let P be the probability measure associated with the reference model (1). An alternative model is defined by a probability measure $P(\eta)$, which is determined by the process η_t for its Radon-Nikodym derivative (likelihood ratio) with respect to P. It is useful to specify models through their Radon-Nikodym derivative since this permits a convenient definition of the set of models that are statistically close to (or difficult to distinguish from) the reference model. I now construct the Radon-Nikodym derivatives under consideration by the agent and describe how they map to specifications of dynamics. The intention is to consider the most general set of dynamics possible, before restricting the alternatives to the subset of models that are statistically close to the reference model.

From the expression for η_t, one can derive the resulting dynamics under $P(\eta)$. Changes to the reference dynamics caused by η are referred to as "perturbations", and the resulting model is called the "perturbed model". Perturbations fall into two categories. The first are perturbations to the diffusion components via changes in the probability law of Z_t. For this category the perturbations considered are completely general, i.e. all equivalent changes of measure are included. The second category are perturbations to the jumps. For tractability, perturbations to the jumps are restricted to changes in the jump intensity and changes to the parameters of the jump size distributions. By Girsanov's theorem for Itô-Lévy Processes we can write $\eta_t = \eta_t^{dZ}\eta_t^J$ where η_t^{dZ} perturbs dZ_t and η_t^J perturbs the jumps.[8]

η_t^{dZ} is defined by the SDE:

$$\frac{d\eta_t^{dZ}}{\eta_t^{dZ}} = h_t^T dZ_t$$

[8]See, for example, Oksendal and Sulem (2007), Theorem 1.31. This multiplicative form arises from the fact that a Brownian motion Z_t and Poisson process N_t defined on the same filtration are independent, which follows from $[Z, N](t) = 0$, i.e. their cross-variation is 0.

where h_t is an n-dimensional process and $\eta_0^{dZ} = 1$. From Girsanov's theorem we have that $Z_t^\eta = Z_t - \int h_t dt$ is a Brownian motion under $P(\eta)$, which implies that $dY_t^c = [\mu(Y_t) + \Sigma(Y_t)h_t]\,dt + \Sigma(Y_t)dZ_t^\eta$. [9] Thus, these perturbations change the drift dynamics and leave the diffusion unchanged. Note that the drift perturbation is driven directly by h_t. Since only the dynamics of $Y_{2,t}$ are ambiguous, I impose $h_t = [0, h_{2,t}]'$, where the 0 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.

η^J is constructed to change the jump intensity and jump size distribution under $P(\eta)$. I discuss the jump dynamics that result for $P(\eta)$ and leave the construction of η^J to Appendix A.1. Consider first the jump intensity. The jump intensity l^η under $P(\eta)$ is given by

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

where a is a scalar parameter. Thus, η^J perturbs the jump intensity by a factor of $\exp(a)$. For the jump sizes, I consider two specific jump size distributions, which are the ones used in the calibration below: (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 shape and scale parameters respectively. η^J is constructed to change the parameters of these distributions so that, under $P(\eta)$, the jump size distributions are:

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

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.[10] Note that, when $\Delta\mu = b = a = 0$ and $s_\sigma = 1$, we are back to the jump distributions of the reference model. Combining the perturbations, the dynamics under $P(\eta)$ can be written as:

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

In addition, denote the moment generating function under $P(\eta)$ by $\psi^\eta(u)$.

[9] The notation Y_t^c means the continuous part of Y_t, i.e. the process obtained by removing the jumps of Y_t.

[10] The shape parameter k is unchanged, though an extension of the setup that allows k to change is possible. Note, however, that changing the scale parameter via b already changes both the mean and variance of the gamma distribution.

Thus, the alternative one-step-ahead dynamics the agent worries about at time t are determined by the set of h_t, a, $\Delta\mu$, s_σ, and b that he considers.[11] The determination of this set is now explained.

3.3 The Size of the Alternative Set

As mentioned earlier, model uncertainty leads the agent to consider a set of alternative models that are statistically close to the reference model – close in the sense that they are difficult to distinguish from the reference model using historical data. 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 alternative model's Radon-Nikodym derivative with respect to the reference model. The agent's set of alternative models is defined by placing an upper bound on the growth rate of alternative models' relative entropy.[12]

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, $\lim_{\Delta t \to 0} \frac{H(t,t+\Delta t)}{\Delta t}$ gives the instantaneous growth rate of relative entropy at time t. It is illustrative to look at this quantity for the diffusion perturbation. A standard calculation (see Appendix A.2) shows that for η_t^{dZ} the instantaneous growth rate of relative entropy is just $\frac{1}{2}h_t'h_t$.

This simple expression says that the rate of relative entropy growth at time t is just half the norm of the h_t vector. Hence, for $h_t = 0$ (the reference model), the rate is 0. As h_t increases, the entropy growth rate increases. This is indicative of the tight link between relative entropy and the 'distance' between $P(\eta)$ and P. Moreover, it indicates that the set of alternative models can be implicitly defined by placing an upper bound on their relative entropy growth rates.

Since $\eta_t = \eta_t^{dZ}\eta_t^J$, the overall relative entropy growth of $P(\eta)$ is the sum of the relative entropy growth rates of η_t^{dZ} and η_t^J. The relative entropy of η_t^J is just the sum of the relative

[11] In other words, this set determines the agent's multiple priors over one-step-ahead probabilities. The agent'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 agent's decision-making is guaranteed to be dynamically consistent.

[12] 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.

entropies for the individual jump perturbations. Appendix A.2 derives the relative entropies for the normal and gamma jump perturbations and gives an expression for the total relative entropy growth rate of $P(\eta)$, which is denoted $R(\eta_t)$. As Appendix A.2 shows, the expression for $R(\eta_t)$ is in terms of $(h_t,\, a,\, \Delta u,\, s_\sigma,\, b)$.

I now exploit the link between entropy and statistical proximity to define the set of alternative models that concerns the agent. The alternative set is defined by choosing all models whose relative entropy growth rate $R(\eta_t)$ is less than some upper bound. The intuition is that, if the relative entropy of a given model is below the bound, then distinguishing this model from the reference model is reasonably difficult. It is then reasonable to be concerned that this alternative model (and not the reference model) is actually the true data generating process. The bound on entropy therefore determines the size of the alternative set. A large bound is interpreted as high uncertainty, since a larger set of models will fall below the bound. In that case, the agent has little confidence in the correctness of his reference model. At the other extreme, a bound of 0 on $R(\eta_t)$ means the alternative set is empty and the agent 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 that controls variation in the level of uncertainty. Hence, the alternative set of dynamics at time t is defined by:

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

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 agent's preferences. If $\varphi = 0$ then the agent has full confidence in the reference model, while increasing the value of φ expands the size of the alternative set to include models that are statistically 'further' away from the reference model. Finally, while φ determines the agent's average level of uncertainty, q_t^2 controls variation in uncertainty over time. When q_t^2 increases, the agent is more uncertain and worries about a larger set of alternative models that includes models that are further away from the reference model.

The mistake of erroneously favoring the reference model over an alternative model that is the true data-generating process, is called making a model detection error. If the alternative model is statistically close to the reference model, the probability of making a detection-error

is high. As the statistical distance between an alternative and the reference model increases, the models become easier to distinguish and the probability of detection error decreases. In calibrating the model, the specific value of φ is chosen to imply a particular probability of making a model detection error under the worst-case model. This is discussed in detail in the calibration section and in Appendix G. The idea is that if the probability of making a model detection error is reasonably high, then the worst-case model is sufficiently difficult to distinguish from the reference model to warrant reasonable concerns about model detection.

3.4 Utility Specification

For a given probability model, the agent's utility over consumption streams is given by the stochastic differential utility of Duffie and Epstein (1992), which is the continuous-time version of the recursive preferences of Epstein and Zin (1989). Denote the agent's value function by J_t and the normalized aggregator of consumption and continuation value in each period by $f(C_s, J_s)$. Therefore, for a given probability model, lifetime utility is given recursively by: $J_t = E_t\left[\int_t^\infty f(C_s, J_s)ds\right]$. The set of probability measures considered by the agent is given by the reference model and alternative set, as described above. The representative agent's utility is then given by:

$$J = \min_{P(\eta)} E_0^\eta \left[\int_0^\infty f(C_s, J_s)ds\right] \tag{4}$$

where E^η denotes expectation taken under the probability measure $P(\eta)$.[13]This utility specifies that the agent expresses his aversion to model uncertainty by being cautious and evaluating his future prospects under the worst-case model within the set of alternatives.

The functional form used for $f(C, J)$ is standard:

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

where δ is the rate of time preference, γ is $1 - \text{RRA}$ (i.e. one minus the agent's relative risk aversion), and $\rho = 1 - \frac{1}{\psi}$, where ψ is the intertemporal elasticity of substitution (IES). An important special case of this aggregator is $\gamma = \rho$, in which case the agent's relative risk

[13]This formulation already embeds the maximization of C_t which in equilibrium is given by the aggregate consumption process.

aversion equals $1/\psi$ and the aggregator reduces to the additive power utility function.

As Epstein and Schneider (2003) show, rectangularity of beliefs implies that J_t solves the following Hamilton-Jacobi-Bellman (HJB) equation:

$$0 = \min_{P(\eta_t)} f(C_s, J_s) + E_t^{\eta}[dJ] \tag{6}$$
$$\text{s.t.} \quad R(\eta_t) \leq \varphi q_t^2$$

The solution of this equation gives the worst-case perturbation, η_t^*, which is needed for asset pricing. Note that in the 'standard' endowment economy framework, one can do pricing by proceeding directly from the Euler equation. Here, the need to solve for η_t^* and the agent's value function adds an extra layer of complexity.[14] Another issue is that there is no guarantee that the worst-case dynamics associated with η_t^* permit tractable asset pricing.[15] In this paper, expressions for the worst-case model, value function, and asset prices are found in closed-form.

4 Solution

I start by expanding the right side of (6):

$$E_t^{\eta}\left[dJ\right] = E_t^{\eta}\left[dJ^c + J_t - J_{t_-}\right] = E_t^{\eta}\left[dJ^c\right] + E_t^{\eta}\left[J(Y_{t_-} + \xi_t \cdot dN_t) - J(Y_{t_-})\right]$$

where J_t^c is the continuous part of J and the second expectation is over the jumps. The Lagrangian corresponding to the minimization in (6) is:

$$f(C_t, J_t)dt + E_t^{\eta}[dJ^c] + E_t^{\eta}\left[J(Y_{t_-} + \xi_t dN_t) - J(Y_{t_-})\right] - \lambda_t\left(\varphi q_t^2 - R(\eta_t)\right) \tag{7}$$

where λ_t is the lagrange multiplier on the (time-t) entropy constraint. We can rewrite the term $E_t^{\eta}[dJ^c]$ by applying Ito's lemma and (2):

$$E_t^{\eta}[dJ^c] = E_t[dJ^c] + h_t^T \Sigma_t^T J_Y dt = E_t[dJ^c] + h_{2,t}^T \Sigma_{2,t}^T J_{Y_2} dt$$

[14] Many papers in related literature have focused on single-state variable or i.i.d environments, which lead to ODEs rather than difficult PDEs. For closed-form solutions, log-utility is often assumed

[15] In i.i.d environments this is not usually a problem because an i.i.d reference model normally leads to an i.i.d worst-case model.

where J_Y is the gradient of J with respect to Y.

Solving for J proceeds as follows. Take first-order conditions with respect to the perturbation parameters $(h_t, a, \Delta u, s_\sigma, b, \ldots)$ and λ_t. Then conjecture and verify a functional form for J that solves the system of first-order conditions and the HJB equation. The solution for J is now discussed while the first-order conditions and other details are left to Appendix D.

4.1 Equilibrium Value Function

Since the aggregator (5) is homogenous of degree γ in the level of consumption and the transition dynamics are independent of the level of consumption, the value function J and HJB equation (6) should also be homogenous of this degree in the level of consumption. This implies the following functional form for the value function:

$$J(Y_t) = \exp\left(\gamma g(\tilde{Y}_t)\right)\frac{C_t^\gamma}{\gamma} = \frac{\exp\left(\gamma g(\tilde{Y}_t) + \gamma \ln C_t\right)}{\gamma} \tag{8}$$

where $g(\tilde{Y}_t)$ is a function of $\tilde{Y}_t$ whose form is not yet specified. Appendix B gives the equation that results from substituting the conjecture into the HJB equation (6). In general, there is no exact analytical solution to that equation. However, I find approximate analytical expressions by approximating a term in this equation. This approximation has been used successfully in the portfolio choice literature (see Campbell, Chacko, Rodriguez, and Viciera (2004)). The approximation log-linearizes the equilibrium consumption-wealth ratio around its (endogenous) unconditional mean. In the case $\psi = 1$ (and any value of γ), the approximation is *exact*, as is the analytical solution. Moreover, as argued in the portfolio choice literature, the approximation is accurate for an interval of values around 1 that easily includes empirical estimates of ψ and the values I use in the calibrations.

The term that is approximated is $\exp\left(-\rho g(\tilde{Y}_t)\right)$. As shown in the Appendix, this is just $1/\delta$ times the equilibrium consumption-wealth ratio. Following Campbell, Chacko, Rodriguez, and Viciera (2004), I log-linearize this term around the unconditional mean of the equilibrium log consumption-wealth ratio

$$\exp\left(-\rho g(\tilde{Y}_t)\right) \approx \kappa_0 + \kappa_1 \rho g(\tilde{Y}_t) \tag{9}$$

where κ_0 and κ_1 are linearization constants whose values are endogenous to the equilibrium solution of the model. The following proposition now provides the solution to the HJB equation.

Proposition 1 *The solution to the HJB equation for $\psi = 1$, or for $\psi \neq 1$ when the log-linear approximation in* (9) *is applied, is given by* (8) *and*

$$g(\tilde{Y}_t) = A_0 + A'\tilde{Y}_t \tag{10}$$

where A_0 is a scalar and A is a vector that gives the loadings on $\tilde{Y}_t$.

The values of A_0 and A are determined by a system of equations. Let $\hat{A} = [1, A]'$ so $\hat{A}$ has the same dimension as Y_t and partition it into $\hat{A} = [\hat{A}_1,\ \hat{A}_2]'$ corresponding to the partition $Y = [Y_1,\ Y_2]'$. Then, A_0, A, κ_0, κ_1, the parameters of the worst-case perturbation (i.e. a, $\Delta\mu$, s_σ, b, ...), and $\tilde{\lambda}$, a constant related to the lagrange multiplier λ_t of the entropy constraint, jointly solve a system of equations that includes the n-dimensional system of equations

$$\begin{aligned} 0 &= \frac{\delta}{\rho}(\kappa_0 + \kappa_1\rho A_0 - 1) + \hat{A}^T\mu \\ 0 &= \tilde{Y}_t^T A\delta\kappa_1 + \tilde{Y}_t^T\tilde{\mathcal{K}}^T\hat{A} - \frac{1}{\tilde{\lambda}}\left(\hat{A}_2^T H_q \hat{A}_2\right) q_t^2 + \frac{1}{2}\gamma\hat{A}^T\Sigma_t\Sigma_t^T\hat{A} + \frac{1}{\gamma}l_1^{\eta'}\left(\psi^\eta(\gamma\hat{A}) - 1\right) q_t^2 \end{aligned} \tag{11}$$

and equations giving κ_0, κ_1, plus a system of equations, defined in Appendix D, that arises from the first-order conditions for the perturbation parameters.

The proof of Proposition 1 is given in Appendix D. A solution of the system of equations given in Proposition 1 and Appendix D verifies the conjecture for the value function. In general, this system of equations must be solved numerically.

Proposition 1 shows that the equilibrium value function is exponential-affine with the vector A giving the elasticities of the value function with respect to the state variables. The sign of an A coefficient determines whether utility rises or falls in the level of the corresponding state variable, and the magnitude of the coefficient measures the impact of that variable on utility. As Proposition 1 notes, the solution to the aforementioned system of equations also gives the solution for the worst-case model, which is now discussed.

4.2 Worst-Case Dynamics

Recall that under alternative models, the perturbation to the drift is $\Sigma_t h_t = [0, \Sigma_{2,t} h_{2,t}]'$. Appendix D shows that under the worst-case model:

$$\Sigma_{2,t} h_{2,t} = -\frac{1}{\tilde{\lambda}} \Sigma_{2,t} \Sigma_{2,t}^T \hat{A}_2 = -\frac{1}{\tilde{\lambda}} H_q \hat{A}_2 q_t^2 \tag{12}$$

where $\tilde{\lambda}$ is a constant that comes out of the equilibrium solution and is closely related to λ_t, the lagrange multiplier on the time-t entropy constraint. The value of $\tilde{\lambda}$ controls the mean size of the perturbation and depends inversely on φ, the average level of uncertainty. A number of additional observations can be made from (12). First, the worst-case drift perturbations are proportional to $\hat{A}_2$, the vector of elasticities of the value function with respect to Y_2 (the ambiguous state variables), but have the opposite sign. This is because if $\hat{A}_2(i) > 0$ then utility rises with $Y_{2,t}(i)$, so the worst-case perturbation should *decrease* the drift of $Y_{2,t}(i)$. In addition, the magnitude of the worst-case perturbation in $dY_{2,t}(i)$ should depend on how important $Y_{2,t}(i)$ is to utility, which is captured by $\hat{A}_2(i)$, in order to allocate entropy in the way that does the most harm to utility. Another observation is that the size of the perturbations vary over time with q_t^2, which controls variation in uncertainty. Hence, when uncertainty is high, the perturbations are large, and vice versa.

Appendix D also derives the equations that determine the worst-case jump perturbations. For jumps the determination of the worst-case perturbations follows the same principle that was just described for the drift: at the minimizing configuration, a given worst-case perturbation optimally trades off the marginal amount of harm it does to utility against its marginal cost in terms of entropy. Thus, the largest perturbations are assigned to aspects of the model where a specification error harms utility in a way that is difficult to detect statistically. The sign of the perturbation to a variable again depends on the sign of its A coefficient. For example, if $\Delta\mu$ is the perturbation to the mean jump in $Y_{2,t}(i)$, and $\hat{A}_2(i) < 0$, then $\Delta\mu < 0$.

Finally, note that, though the reference model was formulated within the affine class, there is no guarantee apriori that asset-pricing under the worst-case dynamics will remain tractable. However, as (12) shows, the perturbation to the drift keeps the worst-case dynamics in the affine class. This is also the case for the jumps. Thus, the worst-case model remains affine, which permits tractable asset pricing.

5 Asset Pricing

Since the representative agent evaluates expectations under the worst-case measure when making his portfolio choice, the Euler equation holds under the worst-case measure. Therefore, assets can be priced using the Euler equation under the worst-case measure. However, we are interested in expected returns under the *reference* model, since it is supposed to be the best estimate of the data generating process based on historical data. While the agent believes that the reference model is the best description of the historical data, he behaves cautiously by pricing assets under the worst-case probabilities. To obtain reference-measure expected returns, expected returns calculated under the worst-case measure are adjusted using (2) to account for the difference in expected dynamics.

5.1 Pricing Kernel

Under the worst-case measure, the pricing kernel is the 'standard' Epstein-Zin kernel. Let M_t denote the time-t pricing kernel. It is convenient to work with the log pricing kernel:

$$d\ln M_t = -\theta\delta dt - \frac{\theta}{\psi} d\ln C_t - (1-\theta) d\ln R_{c,t} \tag{13}$$

where $\theta = \frac{\gamma}{\rho}$ and $\frac{dR_{c,t}}{R_{c,t}} = \frac{dP_{c,t}+C_t}{P_{c,t}}$ is the instantaneous return on the aggregate consumption claim (aggregate wealth). As usual, when $\theta = 1$, (13) reduces to the corresponding expression for CRRA expected utility. To get the log pricing kernel in terms of primitives, we need the return on the consumption claim. Appendix C shows that the consumption-wealth ratio is simply $\delta \exp(-\rho g(\tilde{Y}))$. By market-clearing, the consumption-wealth ratio is also the dividend-price ratio of the aggregate consumption claim. Using this equivalence, the solution for $g(\tilde{Y}_t)$ in (10), and Itô's lemma (with jumps), one obtains[16]:

$$d\ln R_{c,t} = \left[\rho \hat{A}^T + (1-\rho)\delta_c'\right] dY_t + \delta \exp(-\rho A_0 - \rho A' \tilde{Y}_t) dt \tag{14}$$

[16] A useful notational simplification that I use here is: $\rho A' d\tilde{Y}_t + d\ln C_t = \rho \hat{A}^T dY_t + (1-\rho)\delta_c' dY_t$, since $\ln C_t = \delta_c' Y_t$. Rewriting the expression this way makes it possible to collect terms into the single term multiplying dY_t

Note that dY_t includes both the diffusive and jump shocks[17]. Substituting (14) and $d\ln C_t$ into (13) gives the Epstein-Zin (log) pricing kernel:

$$d\ln M_t = -\left[\theta\delta + (1-\theta)\delta\exp(-\rho A_0 - \rho A'\tilde{Y}_t)\right]dt - \Lambda' dY_t \tag{15}$$

where $\Lambda = \left(\frac{\theta}{\psi}\delta_c + (1-\theta)\left[\rho\hat{A} + (1-\rho)\delta_c\right]\right)$. Λ is the vector of risk prices for the economy's shocks. When $\theta = 1$, so that preferences reduce to power utility, $\Lambda = (1-\gamma)\delta_c$, i.e the price of risk on the immediate consumption shock is the agent's RRA and all other risk prices are 0. In general, $\delta_c'\Lambda = 1-\gamma$, i.e. the price of risk for the immediate consumpton shock is the agent's RRA. Recall that (15) is the pricing kernel under the worst-case measure. Therefore, the explicit uncertainty terms do not enter at this point.

5.1.1 The Risk-free Rate

The risk-free rate, $r_{f,t}$, equals $E_t^{\eta}[-\frac{dM_t}{M_t}]$.Since $r_{f,t}$ is known at time t, it is identical under the different measures and no measure adjustment is necessary. Substituting in gives:

$$\begin{aligned} r_{f,t} =& \theta\delta + (1-\theta)\delta\exp\left(-\rho A_0 - \rho A'\tilde{Y}_t\right) + \Lambda^T\left(\mu(Y_t) + \Sigma_t h_t\right)dt \\ &- \frac{1}{2}\Lambda^T\Sigma_t\Sigma_t^T\Lambda dt - l_t^{\eta\prime}(\psi^{\eta}(-\Lambda) - 1) \end{aligned} \tag{16}$$

Uncertainty affects the risk-free rate explicitly through the term $\Lambda^T\Sigma_t h_t$ and via the change in the jump intensity and mgf (jump distribution). It also acts implicitly through the values of A_0 and A. The perturbations decrease expected consumption growth and increase expected variation, which increases the precautionary savings motive. Both effects lower the equilibrium risk-free rate.

5.2 Equity

I derive the return on a share in the stock market. Part of the derivation follows Eraker and Shaliastovich (2008), who derive the market return for an Epstein-Zin representative agent in an affine jump-diffusion setting.

[17] $d\ln R_{c,t} = d\ln R^c_{c,t} + \Delta\ln R$ where $d\ln R^c_{c,t}$ is the continuous part and $\Delta\ln R_{c,t}$ is the jump-related part. Also, $\Delta\ln R_{c,t} = \Delta\ln P_{c,t}$ where $P_{c,t} = \exp(-\ln\delta + \rho g(\tilde{Y}_t))\exp(\ln C_t)$ is the price of the consumption claim.

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. Following Eraker and Shaliastovich (2008), I log-linearize the market return around the unconditional mean of the log price-dividend ratio:

$$d\ln R_{m,t} = \kappa_{0,m}dt + \kappa_{1,m}dv_{m,t} - (1-\kappa_{1,m})v_{m,t}dt + d\ln D_t \tag{17}$$

where $\kappa_{0,m}$ and $\kappa_{1,m}$ are the log-linearization constants. This log-linearization is similar to the one used earlier for the wealth-consumption ratio and represents a continuous-time version of the log-linearization in Campbell and Shiller (1988).

I conjecture that $v_{m,t}$ takes the following functional form:

$$v_{m,t} = A_{0,m} + A_m'Y_t \tag{18}$$

Substituting for $dv_{m,t}$ in (17) gives the log market return in terms of primitives:

$$d\ln R_{m,t} = \kappa_{0,m}dt - (1-\kappa_{1,m})(A_{0,m} + A_m'Y_t)dt + B_r'dY_t \tag{19}$$

where $B_r = (\kappa_{1,m}A_m + \delta_d)$ is the vector of loadings on shocks to the state vector. The return on the market must satisfy the representative agent's Euler equation. Substituting (19) and (15) into the Euler equation and evaluating it under the worst-case measure leads to a system of equations in the unknown coefficients $A_{0,m}$ and A_m. The solution to this system gives the equilibrium values for $A_{0,m}$ and A_m and verifies the conjecture for $v_{m,t}$. Further details are given in Appendix E.

5.3 The Equity Premium

Given A_m, one can find the equity premium. I first give the equity premium under the worst-case measure and then adjust it to get an expression for the equity premium under the reference measure. As usual, the conditional equity premium is given by the covariance of the market return with the pricing kernel. Accounting for the jumps is the only part of this calculation that is not 'standard'. The instantaneous market return is $dR_{m,t}/R_{m,t}$. The

Euler equation implies that $E_t^\eta\left[d(M_t R_{m,t})\right] = 0$.[18] Applying Ito's lemma (with jumps) and substituting in $r_{f,t} = -E_t^\eta[-\frac{dM_t}{M_t}]$ leads to the following expression:

$$\begin{aligned} E_t^\eta\left[\frac{dR_{m,t}}{R_{m,t}}\right] - r_{f,t}\,dt = &-\frac{dM_t^c}{M_t}\frac{dR_{m,t}^c}{R_{m,t}} + E_t^\eta\left[\exp(\Delta \ln M_t) - 1\right] \\ &+ E_t^\eta\left[\exp(\Delta \ln R_{m,t}) - \exp(\Delta \ln M_t + \Delta \ln R_{m,t})\right] \end{aligned}$$

where as before a superscript 'c' refers to the continuous part and Δ refers to the jump part. Substituting into this expression gives:

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

The first term on the right-hand side is (negative one times) the continuous covariance of the market return and pricing kernel. When there are no jumps, this term is the whole equity premium. The remaining term deal with the jumps. This is the equity premium under the worst-case measure. To obtain it under the reference measure, the adjustment $E_t[dR_{m,t}/R_{m,t}] - E_t^\eta[dR_{m,t}/R_{m,t}]$ is added to both sides, giving:

$$\begin{aligned} E_t\left[\frac{dR_{m,t}}{R_{m,t}}\right] - r_{f,t}\,dt = &B_r'\Sigma_t\Sigma_t^T\Lambda - B_r'\Sigma_t h_t \\ &+ l_t{}'\left(\psi(B_r) - 1\right) - \left(l_t^\eta \cdot \psi^\eta(-\Lambda)\right)'\left(\frac{\psi^\eta(-\Lambda + B_r)}{\psi^\eta(-\Lambda)} - 1\right) \end{aligned} \tag{20}$$

where the division in the last term is componentwise. As Appendix F shows, the terms $l_t^\eta \cdot \psi^\eta(-\Lambda)$ and $\psi^\eta(-\Lambda + B_r)/\psi^\eta(-\Lambda)$ actually give the jump intensity vector and moment-generating function vector under the risk-neutral measure. Hence, the second line of the equation can be viewed as the jump-risk premium. The effect of uncertainty in (20) shows up explicitly via the the h_t term and in the perturbed jump intensity l_t^η and moment-generating function ψ^η. Uncertainty amplifies the agent's assessment of the frequency and magnitude of jumps and also makes his assessment of their distribution more pessimistic, thereby increasing the jump risk premia. Note that q_t^2 governs variation in several of the terms in (20), making the level of uncertainty an important driver of the equity premium.

[18] This follows from the condition that $M_t R_{m,t}$ is a η-martingale.

5.4 The Variance Premium and VIX

This section highlights the main reasons that the risk-neutral expectation of return variance (the squared VIX) contains an important uncertainty-related component and how the variance premium is a good filter for this component. Numerical results illustrating these points are provided in the calibration, while complete analytical expressions are derived in Appendix F.

For notational convenience in what follows, let $* \in \{P, \eta, Q\}$ indicate the physical (P), worst-case (η), or risk-neutral (Q) probability measures. I take the reference model as the description of the physical measure. Under the measure $*$ the expectation of integrated return variance from time t to τ is $E_t^* \left[\int_t^\tau (d \ln R_{m,s})^2\right]$, or taking the expectation inside the integral, $\int_t^\tau E_t^* (d \ln R_{m,s})^2$. I want to highlight how this quantity differs across the three measures. Consider $E_t^*(d \ln R_{m,t})^2$, the expectation of the squared return over the first instant. (19) implies that:

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

where $l_t^* = l_1^* \, q_t^2$ and l_1^* is the mean jump intensity under the $*$ measure (i.e. $l_1^P = l_1$, $l_1^\eta = \exp(a) l_t$, and l_1^Q is derived in AppendixF). Note that the jump term depends on the specific measure but the diffusion term does not. Under the worst-case model, the magnitude and frequency of jumps is higher than under the reference model. Thus, $E^P(\xi_t^2) < E^\eta(\xi_t^2)$ and $l_1^P < l_1^\eta$, which implies that:

$$E_t^\eta(d \ln R_{m,t})^2 - E_t^P(d \ln R_{m,t})^2 \;=\; B_r^{2'} \left[E^\eta(\xi_t^2) \cdot l_1^\eta - E^P(\xi_t^2) \cdot l_1^P\right] \, q_t^2 \;>\; 0 \qquad (21)$$

Therefore, over the first instant, expected variance is higher under η than P due to uncertainty about the frequency and magnitude of jumps. Moreover, the difference is a multiple of q_t^2, so it varies directly with the level of uncertainty. In addition, we see that this component has a greater weight in $E_t^\eta(d \ln R_{m,t})^2$ than in $E_t^P(d \ln R_{m,t})^2$ and is therefore reflected more strongly in the variation of worst-case expectations.

Recall that the agent does pricing, which determines risk-neutral probabilities, relative to the worst-case measure. Relative to the worst-case measure, the risk-neutral measure tilts probability mass towards states where marginal utility is high. In particular, states with large, negative jump shocks have greater probability under Q than η, which implies that

$E^{\eta}(\xi_t^2) < E^Q(\xi_t^2)$ and $l_1^{\eta} < l_1^Q$. Combining this effect with (21) gives:

$$E_t^Q(d\ln R_{m,t})^2 - E_t^P(d\ln R_{m,t})^2 \;=\; B_r^{2\prime}\left[E^Q(\xi_t^2)\cdot l_1^Q - E^P(\xi_t^2)\cdot l_1^P\right] q_t^2 \;>\; 0 \qquad (22)$$

To determine the differences in expected integrated variance across measures we also need to know how the quantity $E_t^*\left(d\ln R_{m,u}\right)^2$ evolves as u increases to some later time s. Consider the difference in the 'drift' of this quantity between P and η. Part of the difference is due to the drift perturbation (12) under the worst-case model. This perturbation increases the drift of variables that are harmful to utility. These variables include q_t^2 and other variables driving diffusion volatility, since increased volatility harms utility. As (12) shows, the size of the drift perturbation is a multiple of q_t^2 and is therefore amplified by uncertainty. This means $E_t^*\left(d\ln R_{m,u}\right)^2$ has a higher drift under η than P and by an amount that varies positively with uncertainty. Since we already showed that $E_t^*\left(d\ln R_{m,u}\right)^2$ has a greater initial value under η, this implies that:

$$\int_t^{\tau} E_t^{\eta}\left(d\ln R_{m,s}\right)^2 > \int_t^{\tau} E_t^{P}\left(d\ln R_{m,s}\right)^2$$

Moreover, variation in the difference between these two two quantities covaries with q_t^2.

Moving to the risk-neutral measure further widens the difference in drifts. The combined effect of the worst-case and risk-neutral measure transformations gives that:

$$\int_t^{\tau} E_t^{Q}\left(d\ln R_{m,s}\right)^2 > \int_t^{\tau} E_t^{\eta}\left(d\ln R_{m,s}\right)^2 \qquad (23)$$

The one-month variance premium, $vp_{t,t+1}$, is exactly the difference, for $\tau = 1$, between the two expectations in (23). Thus, we have that $vp_{t,t+1} > 0$. A second point is that $vp_{t,t+1}$ is a good filter for q_t^2 and the level of uncertainty. The reason is that, as (22) indicates, differencing the expectations in (23) (largely) removes the influence of diffusion volatility on $vp_{t,t+1}$. This is an important point if the diffusion volatility is driven by variables other than q_t^2. This makes $vp_{t,t+1}$ a good filter for q_t^2, since it filters out most of the influence of these other volatility drivers and so its variation mostly reflects variation in q_t^2. Since $vp_{t,t+1}$ is part of the risk-neutral expectation of variance (the VIX) we also see that the VIX includes an uncertainty component that is absent from physical expectations of return variance. Finally, recall that the level of uncertainty is an important driver of variation in the equity premium. This implies that equity returns should be predictable by $vp_{t,t+1}$, and to a lesser extent by

the VIX.

6 Calibration

6.1 Reference Model Specification

I now specify the reference model for the calibration. The reference model is an expanded version of the model in Bansal and Yaron (2004) (BY). As in BY, there is a small but persistent component in consumption and dividend growth, which is denoted by x_t. The cash flow processes are given by:

$$\begin{aligned} d\ln C_t &= \left(\mu_c + x_t - \frac{1}{2}\Phi_c^2\sigma_t^2\right)dt + \sigma_t\Phi_c dZ_{c,t} \\ d\ln D_t &= \left(\mu_d + \phi x_t - \frac{1}{2}\Phi_d^2\sigma_t^2\right)dt + \sigma_t\Phi_d dZ_{d,t} \end{aligned}$$

As in BY, ϕ represents the loading of dividend growth on x_t and is greater than 1, reflecting the fact that dividends are much more volatile than consumption.The (conditional) variance of the consumption and dividend growth streams is driven by the stochastic process σ_t^2, which follows an autoregressive process. Hence, σ_t^2 governs the immediate level of *risk* in cash flow growth rates. I assume there is no ambiguity about the structure of these immediate cash flow growth rates and I let $Y_{1,t} = (\ln C_t, \ln D_t)'$. I also make q_t^2 follow an autoregressive process and I let $Y_{2,t} = (\sigma_t^2, x_t, q_t^2)$, so there is uncertainty about the dynamics of the three persistent state variables.

To summarize, the state vector Y_t and transition matrix $\mathcal{K}$ are given by:

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

In addition, let $E(d\ln C_t) = \mu_c$, $E(d\ln D_t) = \mu_d$, $E(x_t) = 0$, $E(\sigma_t^2) = 1$ and $E(q_t^2) = 1$.[19]

[19]The normalization $E(q_t^2) = 1$ was already imposed in Section 3.3.

These values fix the value of the vector μ in the diffusion. Setting $E(q_t^2) = 1$ and $E(\sigma_t^2) = 1$ is a convenient normalization that is without loss of generality. The diffusion covariance matrix is:

$$\Sigma(Y_t)\Sigma(Y_t)' = \begin{bmatrix} H_\sigma \sigma_t^2 & 0 \\ 0 & H_q q_t^2 \end{bmatrix}$$

where $H_\sigma = \text{diag}\left(\Phi_c^2, \Phi_d^2\right)$ and $H_q = \text{diag}\left(\Phi_\sigma^2, \Phi_x^2, \Phi_q^2\right)$. Hence, the diffusions are uncorrelated. Finally, the jump intensity is specified by $l_1 = (0, 0, 0, l_{1,x}, l_{1,q})'$ and the jump sizes are $\xi_t = (0, 0, 0, \xi_{x,t}, \xi_{q,t})'$. The jumps in x_t have a zero-mean normal distribution: $\xi_x \sim N(0, \sigma_x^2)$. The jump sizes in q_t^2 have a gamma distribution: $\xi_q \sim \Gamma(\nu_q, \frac{\mu_q}{\nu_q})$. Specifying a gamma jump size guarantees that the q_t^2 process remains positive. This parametrization of the gamma distribution is convenient since it implies that $E[\xi_q] = \mu_q$. The parameter ν_q is called the 'shape' parameter of the gamma distribution, while the other parameter is the 'scale' parameter. When $\nu_q = 1$, which is the value used in the calibration, the gamma distribution reduces to the exponential distribution.

Several factors motivate the introduction of two volatility processes and the choice of partition of Y_t. First, I wish to separate pure stochastic cash flow volatility from time-varying model uncertainty. The majority of return volatility in structural pricing models comes from cash flow volatility, as is also the case below. However, it need not be the case that uncertainty moves in lock-step with cash flow (or return) volatility and creating separate volatility processes enables the model to capture this potential separation. In terms of the partition of Y_t, it is reasonable that model uncertainty should be much less important for immediate cash flows than for the dynamics of the state variables. The immediate cash flow growth rates are comparatively easy to observe and measure, and they have low persistence. On the other hand, the state variable dynamics are hard to measure and are potentially quite persistent. The persistence means that relatively small, difficult-to-detect perturbations to the state variable dynamics may have large cumulative effects. Model uncertainty is then particularly relevant for the dynamics of these variables. Moreover, since shocks to these persistent variables can have large effects, it is reasonable that the level of uncertainty and the risk of these shocks move together. In Section (6.4.1) I also present results for a model where the level of uncertainty is independent of all cash flow risks in order to illustrate in a stark way the pure effects of time-varying uncertainty.

6.2 Parameter Values

In calibrating the model I use the following guidelines. I aim to find parameter values for the model specification such that (i) once they are time-averaged to an annual level, the model's 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 (iii) the model matches moments of market return volatility, the VIX, and the variance premium, as well as the projections of 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. For the model calibration, I normalize the parameter values to a monthly interval, i.e. $\Delta t = 1$ is one month. Table IV provides the parameter values for the calibration, which are now discussed.

The monthly normalization makes it easy to compare the parameter values to those in Bansal and Yaron (2004). The cash flow parameters are similar to those in Bansal and Yaron (2004), though x_t here is somewhat less persistent. 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. Hence, the value of ρ_x in Table IV indeed implies that x_t has high persistence and represents a long-run component in consumption and dividend growth. As in BY, ϕ represents the sensitivity of dividend growth to the long-run component, which is greater than that of consumption growth. The volatility and uncertainty processes, σ_t^2 and q_t^2, are also persistent, though significantly less so than the volatility process in BY.

Table IV also includes the jump 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 1 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'. The Table also shows that 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 instrumental in capturing the high variance premium (and high price of options). Note that the frequency of these jumps is comparable to what has been found in studies on the empirical properties of returns and in empirical option-pricing literature (see e.g. Singleton (2006)).

Finally, the table shows the preference parameters. Relative risk aversion is set to 5,

which is right in the middle of the range considered by Mehra and Prescott (1985), and is far lower than the levels of risk aversion typically needed to match the equity premium. The agent's aversion to model uncertainty is an important part of the reason that this low risk aversion is able to match the equity premium. The IES is set 2, which corresponds to the estimate from Bansal, Kiku, and Yaron (2007) and dampens the level and volatility of the risk-free rate.

Finally, the mean level of uncertainty is set by the value of φ. This parameter's value is not interpretable directly. Rather, it is mapped to the detection error probability for the worst-case model. Detection error probabilities are a useful tool for calibrating model uncertainty that is due to Anderson, Hansen, and Sargent (2003). Suppose the worst-case model is the true data-generating process. The detection error probability is the probability that a likelihood ratio test will mistakenly reject the worst-case model in favor of the reference model based on a sample of data of some finite length. It is given by the probability that, under the worst-case model, the log-likelihood ratio of the worst-case and reference models is negative. The parameter φ and the degree of uncertainty are deemed reasonable if the corresponding worst-case model is sufficiently similar to the reference model that the detection error probability is sufficiently high. If the detection error probability is sufficiently high, then this means it is difficult to reject the worst-case model and the parameter φ is deemed reasonable. Appendix G explains how to calculate exact detection error probabilities using the Radon-Nikodym derivative of the worst-case model. For simple i.i.d environments there is an analytical formula. For the calibrated model, this calculation requires solving a set of ODEs for the characteristic function of the Radon-Nikodym derivative and then calculating a Fourier inversion numerically.[20] The detection error probability depends on the length of the data history the agent considers for inference regarding the data-generating process. This history is limited by both the availability of data records and also by any structural breaks that may have occurred within the sample. Anderson, Hansen, and Sargent (2003) consider 10% as a reasonable bound on the detection error probability. My choice of φ corresponds to an 11% detection error probability for the post-war sample. Note that two indistinguishable models correspond to a 50% (not 100%) detection error probability. Naturally, if the sample is shorter then the detection error probability is higher. For comparison, the same φ implies detection error probabilities of 16.4%, 26.0%, and 47.5% for sample lengths of 40 years, 20 years, and 1 year respectively. These are non-trivial chances of detection error, so concerns

[20]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 Fourier methods.

about incorrectly rejecting the worst-case model seem quite reasonable. It is interesting to note that infrequent jumps can lead to higher detection error probabilities. This is clear if one thinks of the extreme case of rare-events or 'peso problems', where detection errors are likely even with very long samples of data.

6.3 Results

Table V provides the empirical moments and the corresponding statistics for the calibrated model. In order 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 1000 simulations, each based on the same sample length as its data counterpart. The time increment used in the simulations is one month. For the consumption and dividend dynamics I utilize the longest sample available, (1930:2006), so the the simulations are based on 924 monthly observations which are time-averaged to an annual sample of length 77, 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 is monthly and available only for the latter part of the sample (1990.1-2007.3). Thus, the model's variance premium-related statistics are based on the last 207 monthly observations in each of the 1000 simulations. It is important to note that the *reference* model's dynamics are the ones being simulated. 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 agents and that it provides a good fit to the historical data (or in fact generated it). Under this view, the data point estimates should be within the 5%-95% percentile intervals generated by the model simulations. For completeness, I also provide HAC robust standard errors for the data statistics.

The top panel in Table V shows that the reference model captures quite well several key moments of annualized consumption and dividend growth. The data-based mean and volatility of dividends and consumption growth are in fact close to the median estimates from the model and fall well within the 90% confidence interval. It is important to note that even with the presence of jumps, the distribution of model moments is very reasonable. 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 agents to use as their reference model.

The second panel of Table V presents some of the model's asset pricing implications. This panel shows annual data on the market, risk free rate and price-divided ratio. As mentioned earlier, the corresponding model statistics are time averaged annual figures. The panel shows that the model does a good job in capturing the equity premium and the volatility of excess returns. The model is able to match the equity premium even with a relative risk aversion of only 5. This reinforces the conclusion of other equilibrium models that have included robustness concerns, for example Maenhout (2004) and Liu, Pan, and Wang (2005). However, unlike the models in these two papers, the model here matches the equity premium while also matching the properties of the consumption and dividend processes. Moreover, the model here is able to simultaneously account for the volatility of the market return. The table further shows that the model captures the low mean and volatility of the risk free rate. The rows labeled 'skew' and 'kurt' give the skewness and kurtosis of *monthly* excess returns for the sample (1930:2006). The point of including these moments is to show that the dynamics of the model, particulary the jumps, do not cause the return distribution to be excessively heavy-tailed. Moreover, they show that the model *does* capture the negative skewness and high kurtosis observed in the data.[21] The one moment where the model falls somewhat short is in generating the large volatility of the price-dividend ratio.

The bottom panel in Table V provides a number of statistics pertaining to integrated variance and the variance premium, all at the monthly horizon. The impact of time-varying model uncertainty shows up very strongly here. Note first that the model is able to generate a large average variance premium ($E[VP]$), so that the model-based risk-neutral expectations of variance are substantially larger on average than true expectations. Furthermore, the model's conditional variance premium is also volatile, as in the data. Both results reflect the substantial impact of time-varying uncertainty concerns, as discussed in Section 5.4. The table further shows that the model's median skewness and kurtosis for the variance premium are right in line with the large values in the data, which is a result of periodic spikes in the level of uncertainty. The first two lines of the panel also show that the model almost exactly matches the volatility of expected integrated variance under the two measures [$\sigma(\text{var}_t^P(r_m))$ and $\sigma(\text{var}_t^Q(r_m))$], and that the risk-neutral expectation of integrated variance varies substantially more than the physical expectation, as in the data. Finally, the model's median autocorrelation for the P and Q expectations of integrated variance are close to

[21] The kurtosis estimate may appear high relative to some other estimates. This is due to starting the sample in 1930. By comparison, the (1950:2006) estimate for the kurtosis of monthly excess returns is 6.00 (1.74) and skewness is -0.78 (0.35).

the data and are easily within the 90% confidence interval. This shows that (monthly) conditional return volatility inside the model is persistent, but not extremely so, as in the data.

There are two main ways that uncertainty helps to produce a large and volatile variance premium. As discussed in Section 5.4, several of the perturbations in the worst-case model cause an increase in variance. Both the jump intensities and magnitudes are increased under the worst-case model. For the calibrated model, approximately 57% of the entropy is 'allocated' to the jump perturbations. In addition, shocks to the drift perturbation, which scales with the level of uncertainty, also increase expected variation. Time-variation in uncertainty also leads to a large variance premium because it causes shocks to q_t^2 to carry a large, negative price of risk, which further increases risk-neutral expectations of integrated variance. One reason q_t^2 shocks carry a high risk price is that q_t^2 drives variation in the level of uncertainty and increases in uncertainty adversely affect the agent's utility. In addition, the drift perturbation in q_t^2 under the worst-case model causes its autocorrelation to increase relative to the reference model, which increases the perceived impact of of any shocks to q_t^2. Therefore, the agent wants to hedge increases in q_t^2 and is willing to pay a high premium for assets, such as options, that have a high payoff when uncertainty spikes up. Section 6.4.1 further analyzes the impact of time variation in uncertainty.

An increase in uncertainty and loss of confidence in the reference model increases the distance between the worst-case and reference models. The agent then perceives that growth prospects are worse (through the perturbations to x_t) and that jump risks are higher. Both affects cause prices to decrease and increase expected returns. High levels of uncertainty are therefore associated with high expected equity returns and the level of uncertainty has predictive power for stock returns. Since the variance premium acts as a filter for the level of uncertainty, it should be a predictor of excess stock returns. Section 2 highlighted the predictive ability of the variance premium in the data. As the bottom panel of Table V shows, the model captures this predictability. The bottom of the panel presents the projection coefficients in the data and in the model for predictive regressions of 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 close to 2% and the 90% finite sample distribution of R^2 clearly includes the 1.5% R^2 from the data. For the 3

and 6-month ahead projections, the median R^2 increases to 4.0% and 5.4% inside the model, which, though high, is similar to the 5.9% and 4.0% values in the data. Overall, the results of Table V indicate that the model can capture quite well the cash flow, asset pricing and variance-related moments in the data.

Finally, one may wonder how the 5%-95% sampling intervals of the simulated model moments compare with the standard errors for their empirical estimates? For the top panel of cash-flow moments, the Monte Carlo sampling distributions are roughly symmetric and have widths consistent with the asymptotic standard errors. This is also generally the case for the middle panel of return moments. However, for some of the variance-related moments in the bottom panel, the sampling intervals are clearly much wider their corresponding asymptotic distributions. Since these are moments of heavy-tailed distributions and the data samples are relatively short, their 5%-95% sampling intervals may be poorly described by the HAC estimates of standard errors. For comparison, I compute HAC standard errors for the moment estimates within the model simulations. Table VI shows that even though the sampling intervals for the variance moments are wide, the median standard errors are actually quite close to their empirical counterparts.

6.3.1 Variance under P and Q and Predictability

There is a long literature in empirical asset pricing that has looked at whether the conditional variance of stock returns predicts expected excess stock returns. This literature has come to mixed conclusions, with some early studies finding predictive power while others have not. The recent work in Ghysels, Santa-Clara, and Valkanov (2005) claims to find this relation by more precisely measuring conditional volatility. Bollerslev and Zhou (2007) first pointed out that the variance premium appears to be a stronger predictor of excess stock returns than conditional variance. The model in this paper makes predictions that are consistent with this finding. Within the model, both σ_t^2 and q_t^2 drive variation in the conditional variance of returns. However, the majority of the variation in the equity risk premium comes from variation in q_t^2. As pointed out in Section 5.4, the variance premium is a essentially a filter for q_t^2, and should therefore be a better predictor of excess returns than just the conditional variance.

This point can be seen in Table VII, which provides a more in-depth look at the properties of expected integrated variance under the P and Q measures. The top panel displays prop-

erties of P-measure expected integrated variance, which corresponds closely to conditional one-month variance. The bottom panel looks at the Q-measure expectation, which replicates the (squared) VIX inside the model. Comparing the P-measure variance's predictive R^2 for excess stock returns to the predictive power of the variance premium in Table V, we see that inside the model the variance premium has stronger predictive power than the conditional variance. At the median values for both the 1 and 3 month horizons, the predictive power of the variance premium is over 50% greater than for the P-measure expected variance. The bottom panel shows that the predictive power of Q-measure expected variance (i.e the VIX^2) falls in between that of conditional variance and the variance premium, for the two maturities both in the data and the model. As explained earlier, this results from the fact that the risk-neutral expectations more strongly reflect the level of q_t^2, which is an important driver of the conditional equity premium. This is further indicated by the skewness and kurtosis statistics of both the model and data, which reflect the fact that variation in q_t^2 has the greatest influence on the variance premium and the smallest influence on P-measure expectations of variance. Note that the model matches the empirical skewness and kurtosis moments very well. Along the same lines, the table also shows that the median correlation in the model between the variance premium and Q-measure expected variance is higher than the correlation between the variance premium and P-measure conditional variance. This is consistent with the data. The model also does a very good job matching the actual correlation values and the fact that the correlation between any two of the three series is quite high.

6.3.2 Option Prices and the Volatility Surface

Although the variance premium is a statistic that captures the premia embedded in option prices, one may also be be interested in the whole volatility surface implied by the model. To that end, I determine option prices for the calibrated model and calculate their Black-Scholes implied-volatilities. I calculate the option prices by using Fourier-transform techniques. I first solve a system of ODEs for the characteristic function of the state vector as in Duffie, Pan, and Singleton (2000) and then calculate and invert the option transform based on the method of Carr and Madan (1999). Figure 1 contains two plots that show the model-based and empirical implied volatility curves at maturities of 1, 3 and 12 months 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 curves in the top plot are the average daily implied

volatilities for options on the S&P 500 index traded in the over-the-counter market. The data is obtained from Citigroup and covers October 1999 to June 2008. The range of strikes available in the OTC market for index options is often much broader than for exchange-traded options. See Foresi and Wu (2005) for more details on the OTC market. The model-based option prices are computed by setting the state vector Y_t equal to its unconditional mean.

Figure 1 shows the the model does quite a good job matching both the overall shape and actual values of the implied volatility curves. In particular, it captures the vol skew, the steep slope in implied vols for out-of-the-money put prices (low moneyness). This feature exists at all three maturities, but is particularly pronounced for the 1-month maturity. The model also does a very nice job replicating the decay in the skew as the horizon increases. It further matches the shallow positive slope in the vol surface for at-the-money vols, whereby the 1-year at-the-money vol is the highest and the 1-month vol the lowest.

Figure 2 provides a closer look at the one and three month implied vol curves. The top plot shows the 1-month curves. Note that in addition to doing a good job in matching the very steep slope in the 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 happens for the 3-month curves in the bottom plot. The model fit is very good for low moneyness strikes but not as good for high-moneyness.

6.3.3 Implied Time Series of State Variables

I use the model-based option prices to extract from the empirical options data an implied time series of q_t^2 and σ_t^2. I do this by finding the values of these two state variables that provide the best fit to a cross-section of implied-volatilities at each date t under the calibrated model. Although technically x_t also effects option prices via the risk-free rate, this effect is minuscule and is should be swamped by any noise in the price data. I therefore fix x_t at its unconditional mean. Thus, using at least two implied-volatilities on a given date t, one can solve for the values of the two state variables that generates the best model fit to the data. While using exactly two implied-volatilities generates an exact fit, one can include additional options and do a (non-linear) least-squares fit in order to attenuate the impact of measurement error.

Figure 3 shows the extracted series implied by fitting model-based implied volatilities their empirical counterparts for strikes with moneyness of 1, 0.9, and 0.8 at a 1-month maturity. A moneyness of 0.8 represents put options that are far out-of-the-money and should be informative regarding variation in investor's fears of negative jump shocks. The figure shows that the implied q_t^2 series is volatile and is occasionally hit by extreme spikes, while the implied σ_t^2 series is much smoother. The extreme spikes in q_t^2 correspond to the the periods of the 1998 LTCM crisis, September 11th, and the corporate scandals of 2002. Both series are low and tranquil for a period that begins in 2004 and ends in 2007. Since the extracted series were not used directly in the calibrated model, a question is whether their time-series properties are consistent with the properties of the calibrated model. This appears to be largely the case. The means of the implied $(\sigma_t^2,\, q_t^2)$ are $(1.17, 1.04)$, which is close to their population means of 1, while their standard deviations are $(0.64, 1.12)$, which is close to the population values of $(0.69, 1.19)$. The autocorrelations of the implied (σ_t^2, q_t^2) are 0.80 and 0.68 respectively, which are somewhat lower than their population counterparts, while the correlation between the two series is 0.15.

6.4 The Impact of Uncertainty

Table VIII conducts a two-part comparative statistics exercise on the model of Table IV 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 with regards to only the jump shocks in the model, leaving on uncertainty regarding the diffusive parts of the dynamics. Thus, the jump parameters do not change in moving from the reference to the worst-case model. The second panel, Model 1-B, turns off all model uncertainty, so the agent has full confidence in the reference model. This exercise is intended to assess the impact of model ambiguity on the asset prices. The table shows only the asset price data, since the reference model cash flow dynamics are unaffected.

The top panel, Model 1-A, shows that eliminating ambiguity with respect to the jump shocks reduces the median equity premium, though the equity premium remains nontrivial. The other moments in the top panel are not greatly affected. Return volatility is reduced a little and the kurtosis of the monthly returns decreases somewhat. Note that for model 1-A the 'distance' to the worst-case model is the same as before (i.e. the relative entropy of the worst-case model is the same). For Model 1-B, where entropy is zero and the worst-case

model is reduced to the reference model, the median equity premium becomes quite small.

The bottom panel shows the impact of uncertainty on the variance premium. The results for Model 1-A show that eliminating uncertainty regarding the jump components of the model greatly reduces the size and volatility of the variance premium. The average variance premium is reduced by almost an order of magnitude and the 90th percentile of the simulations is nowhere near the data estimates. Eliminating all ambiguity in Model 1-B reduces the average variance premium even further, so that it is essentially zero. Moreover, the predictive R^2 of the variance premium is reduced successively in the two models. As pointed out by Drechsler and Yaron (2008), the variance premium's predictive power comes largely from the fact that it reveals the probability (intensity) of jump shocks. When the jump intensity is not directly amplified under the worst-case model, as in Model 1-A, the jumps' importance decreases along with the predictive power of the variance premium. Without any model uncertainty, as in Model 1-B, the influence of q_t^2 on risk premia is greatly diminished, and so is its predictive power. This is apparent in the diminished median R^2s shown in the table. Lastly, as the variance premium becomes very small, the predictive regression coefficients become very unstable. If the variance premium was this small empirically, it would likely be obscured by estimation noise.

The results from the table are reinforced by Figures 4 and 5, 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, though it is just as much a problem at maturities of 3 and 12 months.

Finally, I consider shutting off all model uncertainty (as for Model 1-B) and then raising the risk aversion to the point where the equity premium matches that of the benchmark calibration. The risk aversion required for this turns out to be 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 corresponding values under the benchmark calibration. Thus, in this case a high risk aversion alone is insufficient for jointly capturing the variance and equity premia. Moreover, a risk aversion of 13.3 exceeds the value of 10 which is often viewed as the upper end of the reasonable range for risk aversion, and is substantially higher than the risk aversion of 5 in the benchmark calibration. Hence, for the benchmark calibration, model uncertainty, especially with regards to the jump shocks, is an important component for reconciling the

asset pricing moments with the properties of fundamentals and a reasonable level of risk aversion.

6.4.1 Illustrating the Impact of Time-Variation

To focus on the mechanism by which time-variation in uncertainty contributes to the level and variation of the variance premium, I consider a modified version of the main model from the calibration. Starting with the main model, I set x_t to be homoscedastic and I turn off jump shocks in x_t. Thus, the gaussian volatility of x_t is Φ_x rather than $\Phi_x \times q_t$ and $l_{1,x} = 0$ (no jumps). The rest of the reference model dynamics are unchanged. Notably, I leave in jump shocks in q_t^2. I restrict uncertainty solely to x_t, so that the only perturbation under the worst-case model is in the drift of x_t. This drift perturbation takes the same form as in (2) and is given by the scalar $-\frac{1}{\tilde{\lambda}}\Phi_x^2 A(x) q_t^2$, where $A(x)$ is the loading of the value function on x_t (the x_t component of the vector A, see Proposition 1). Since utility loads positively on x_t, $A(x) > 0$, the perturbation decreases the drift in x_t under the worst-case model. This is intuitive as the agent fears lower consumption growth rates. As before, the perturbation scales with q_t^2 so that increases in uncertainty result in a more negative assessment of the drift in growth rates.[22]

Unlike the main calibration model, where the uncertainty process co-moves with the conditional volatility in x_t, here x_t is homoscedastic and q_t^2 is independent of the rest of the processes. Although assuming that q_t^2 is orthogonal to the other processes seems extreme, it allows me to starkly illustrate some mechanisms inherent in the main model. I illustrate that time-variation in uncertainty helps substantially to generate: (1) the level of the variance premium (2) variation in the variance premium and the VIX (3) return predictability by the option-related measures. As is the case for the main model, the inclusion of sharp increases in uncertainty (jumps) has a big impact on the magnitudes of these effects. To see how time-variation in uncertainty contributes to the variance premium in this model, consider the effect of an upward spike in uncertainty. The result is a shock downward in the agent's worst-case assessment of the drift in the long-run growth rate. Since the agent perceives lower future growth rates, the equity return in such states tends to be large and negative. Moreover, such uncertainty shocks represent bad states of the world and therefore carry high

[22]To maintain the affine structure for the perturbation I change the structure of the entropy constraint to $R(\eta_t) \leq \varphi(q_t^2)^2$. This effects calculation of the detection error probabilities. As this model is used for illustration, I do not discuss the exact calculation of the detection error probabilities.

state-prices. The combination of potentially large return moves and high state-prices implies a high risk-neutral conditional variance and a substantial variance premium.

To illustrate this quantitatively, I simulate the model as well as a variant of it that fixes the level of uncertainty and compare their resulting statistics. The caption in Table IX gives the parameters used in the simulation. The major difference relative to the main model calibration is that I amplify the intensity and size of the jumps in q_t^2. This is done for illustrative purposes, as the model is missing several sources of risk relative to the main calibration. For the variant with constant uncertainty, I turn off all variation in q_t^2 and raise risk aversion to get the same equity premium. Table IX shows the simulation outputs. The left side of the table gives the results with time-varying uncertainty and the right side is for constant uncertainty. Notice that the model with time-varying uncertainty produces a modest but non-trivial variance premium level and volatility. In contrast, the right side shows that shutting off time-variation in uncertainty drops the variance premium to essentially 0. Similarly, with time-varying uncertainty, the predictive power of the variance premium for one and three month returns is sizeable at around 1.6% and 3.5% but is much lower when time-varying uncertainty is shut off. In addition, with time-varying uncertainty the predictive power of the variance premium is substantially larger than that of the conditional variance, with the VIX's falling in between. This result, which was discussed for the main model calibration, arises because the variance premium acts like a filter for q_t^2, so that variation in q_t^2 is reflected more strongly in the variance premium than in the conditional variance. The contrast in predictive power is more stark here than for the main model since here q_t^2 is unrelated to volatility. On the right side, by construction, the predictive power of all three quantities is identical, since there is only one driver of variation (i.e. σ_t^2). Similarly, the left side shows that the conditional variance and VIX are more volatile when uncertainty is time-varying. This is not surprising as variation in uncertainty represents a second channel causing changes in the conditional variance.

6.5 Forecast Dispersion and the VIX

Figure 6 provides further support for the link between the VIX and (model) uncertainty. To measure uncertainty I use the dispersion in forecasts of next quarter's real GDP growth from the Philadelphia Fed's Survey of Professional Forecasters (SPF). The dispersion is measured simply as the standard deviation in the growth forecasts. Since the forecasters have access

to the same public information, the dispersion in their forecasts should capture differences in their models.[23] If we take the set of models they use as the representative agent's alternative set of models, then the dispersion should capture the size of the alternative set. This approach to measuring uncertainty is closely related to the empirical approach in Anderson, Ghysels, and Juergens (2007).

The figure plots the quarterly dispersion measure along with the value of the VIX at the end of the previous quarter. The two series appear to be strongly related. Their correlation is 0.48, with a standard error of 0.11. Moreover, the two series tend to spike at the same time, particularly in 1987-88, 1990-91, and 2001-2002. The economic turmoil of 2008 has also caused both to spike. One notable exception to their strong comovement is the financial crisis of 1998, which caused a sharp spike in the VIX without a corresponding strong increase in forecast dispersion. Perhaps this exception can be attributed to the fact that the 1998 episode was quite short-lived and also did not appear to be directly related to economic events in the US. For completeness, I also calculate the correlation between the level of q_t^2 and the model-implied VIX for the main calibration. For the calibrated model simulations, the 5, 50, and 95 percentiles of this correlation are 51%, 78%, and 93% respectively. The high correlation confirms that the model-generated VIX strongly reflects uncertainty, though it is also clear that the correlation is not perfect.

7 Conclusion

An important aspect of studying asset prices in equilibrium models is that fundamental risk prices arise endogenously from the solution of the model and depend jointly on dynamics and preferences. This can make it difficult to match empirically observed risk premia, such as the large variance premium embedded in option prices. Time-varying model uncertainty presents an intuitively appealing and promising direction for explaining these high risk premia. To explore this avenue quantitatively requires building models with a structure rich enough to enable uncertainty concerns to have their full implications. Tractability is an obstacle to solving for equilibrium in these models and imposes an additional layer of complexity in solving for asset prices. This paper builds a flexible equilibrium framework with time-varying model uncertainty concerns and solves for the 'worst-case' model and equilibrium

[23] The Philadelphia Fed makes a point of sending with their questionnaire the data from the BEA's advance report and other recent reports of economic data.

asset prices.

The paper conducts a quantitative calibration of a model where there is time-varying uncertainty regarding the diffusive and jump shock shocks to the mean and volatility of long-run cash-flow growth rates. Jump shocks and persistent growth dynamics present important model specification concerns since perturbations to them are potentially harmful to utility and are also statistically difficult to detect. The calibrated model is able to generate the large variance premium in options and the high vol skew while simultaneously matching the moments of cash-flows, the equity premium and risk-free rate. It demonstrates that uncertainty and its time-variation are central to capturing the variance premium and the ability of option-related quantities to predict equity returns. It also predicts, consistent with the data, that these option-implied measures should have greater predictive power for returns than statistical variance measures, since they reflect the level of uncertainty more strongly.

Appendix

A Measure Change and Entropy for Jumps

A.1 Derivation of Measure Changes

Recall that $\eta_t = \eta_t^{dZ}\eta_t^J$. I derive expressions for η_t^{dZ} and η_t^J corresponding to the alternative model dynamics discussed in the main text.

η_t^{dZ} solves the SDE $\frac{d\eta_t^{dZ}}{\eta_t^{dZ}} = h_t^T dZ_t$. An application of Ito's lemma shows that its solution is:

$$\eta_t^{dZ} = \exp\left(\int_0^t h_s^T dZ_s - \frac{1}{2}\int_0^t h_s^T h_s ds\right)$$

Note that η_t^{dZ} is a martingale and that $\eta_0^{dZ} = 1$. Let $P(\eta)$ be the measure that results from application of η to the reference measure P. Girsanov's theorem then implies that $Z_t^\eta = Z_t - \int h_t dt$ is a Brownian motion under $P(\eta)$. Writing the dynamics (1) in terms of Z_t^η alters the drift by adding to it the term $\Sigma(Y_t)h_t$, as in (2). This accounts for the perturbation to the drift under $P(\eta)$.

Since the Poisson process arrivals are (conditionally) independent and the jump sizes are i.i.d, the expression for η_t^J can be written as $\eta_t^{J_1}\eta_t^{J_2}\ldots$ where $\eta_t^{J_i}$ changes the probability law for the i-th jump component. I construct such terms to change the distribution of gamma-distributed jumps and normally distributed jumps.

Consider first gamma-distributed jumps, $\xi_i \sim \Gamma(k,\theta)$, where k and θ are the shape and scale parameters respectively. I want to construct the corresponding term $\eta_t^{J_i}$ in the Radon-Nikodym derivative so that under $P(\eta)$ the jump distribution is given by $\xi_i^\eta \sim \Gamma\left(k, \frac{\theta}{1-\theta b}\right)$, where b is the parameter that changes the gamma distribution's scale. I further specify the measure change so that the corresponding jump intensity changes from $l_{t,i}$ to $l_{t,i}^\eta = \exp(a)l_{t,i}$, i.e. it is scaled by the term $\exp(a)$ where a is a perturbation parameter. The desired $\eta_t^{J_i}$ solves the following SDE:

$$d\eta_t^{J_i} = \left(\exp\left[a + b\xi_i - \ln\psi_i(b)\right] - 1\right)\eta_{t-}^{J_i} dN_t - \left(\exp(a) - 1\right) l_{t,i}\eta_t^{J_i} dt$$

where $\psi_i(b)$ is the moment-generating function of ξ_i evaluated at b. An application of Ito's lemma shows that $\eta_t^{J_i}$ is given by:

$$\eta_t^{J_i} = \exp\left(\int_0^t \left(a + b\xi_{i,s} - \ln\psi_i(b)\right) dN_s - \int_0^t l_{s,i}\left(\exp(a) - 1\right) ds\right)$$

Note that the process $\eta_t^{J_i}$ is a martingale and $\eta_0^{J_i} = 1$. Girsanov's theorem for jump processes

then implies that under $P(\eta)$, the jump intensity is scaled by $\exp(a)$, as desired. Furthermore, under $P(\eta)$ the moment-generating function of ξ_i is given by:

$$\psi_i^\eta(u) = \frac{\psi_i^\eta(b+u)}{\psi_i(b)}$$

Straightforward substitution of the mgf for a gamma distribution shows that $\psi_i^\eta(u)$ is the mgf of a $\Gamma\left(k, \frac{\theta}{1-\theta b}\right)$, as desired.

Finally, I consider normally-distributed jumps, $\xi_k \sim \mathcal{N}(\mu, \sigma^2)$. I want to construct the corresponding term $\eta_t^{J_k}$ in the Radon-Nikodym derivative so that under $P(\eta)$ the jump distribution is given by $\xi_k^\eta \sim \mathcal{N}\left(\mu + \Delta\mu, \sigma^2 s_\sigma\right)$, where $\Delta\mu$ shifts the mean and s_σ scales the variance of the distribution. I further specify the measure change so that the corresponding jump intensity changes to $l_{t,k}^\eta = \exp(a) l_{t,k}$. The desired $\eta_t^{J_k}$ is given by:

$$\eta_t^{J_k} = \exp\left(\int_0^t \left(a + b_2\xi_k^2 + b_1\xi_k - \frac{1}{2}\left[\frac{(\mu+\Delta\mu)^2}{s_\sigma\sigma^2} - \frac{\mu}{\sigma^2} + \ln s_\sigma\right]\right) dN_s - \int_0^t l_{t,k}\left(\exp(a) - 1\right) ds\right)$$

where $b_1 = \frac{\mu(1-s_\sigma)+\Delta\mu}{s_\sigma\sigma^2}$ and $b_2 = \frac{1}{2}\frac{1}{\sigma^2}\left(1 - \frac{1}{s_\sigma}\right)$. By construction, the process $\eta_t^{J_k}$ is a martingale and $\eta_t^{J_k} = 1$.

Finally, since the the terms composing η_t are all martingales and have zero cross-variation (they are conditionally independent) η_t is a martingale with $\eta_0 = 1$ and therefore the measure $P(\eta)$ is indeed a probability measure.

A.2 Derivation of Relative Entropy Growth for Jumps

As discussed in the main text, the relative entropy growth rate is defined as $R(\eta_t) = \frac{d}{ds}|_{s=0} E_t^\eta\left[\ln \eta_{t+s}\right]$. 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 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 A.1). 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 0. It is then 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 Appendix A.1). To find $R(\eta_t^{J_i})$, I recall two facts: (1)

$l_t^\eta = \exp(a) l_t$ (2) $E_t^\eta[\xi_i] = \frac{k\theta}{1-\theta b}$. Using these facts, a straightforward calculation shows that:

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

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 Appendix A.1). Straightforward calculations and algebraic simplification show that:

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

B HJB Equation Derivation

Substituting (8) into the aggregator (5) gives:

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

Using Ito's lemma, we have that:

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

where **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 HJB equation (6) can then be rewritten as:

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

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 $Y_{2,t}$.

C Equilibrium Consumption-Wealth Ratio

In equilibrium, markets clear so that the representative agent must hold all of his wealth in the aggregate consumption claim. To derive the equilibrium consumption-wealth ratio, consider the consumption and portfolio problem of the representative agent in this endowment setting. Under the reference measure, the price of the aggregate consumption claim P_c follows an Itô process of the form:

$$dP_{c,t} = (P_{c,t}u_{c,t} - C_t)dt + P_{c,t}\sigma_{c,t}^T dZ_t + P_{c,t-}(\exp(\Delta \ln P_{c,t}) - 1)$$

There is also a risk-free money market account in zero-net supply, paying an endogenously determined rate $r_{f,t}$. The agent chooses the proportion α_t of his wealth, W_t, to invest in the consumption claim. His budget constraint is then:

$$dW_t = W_t \left[\alpha_t(u_{c,t} - r_{f,t}) + r_{f,t}\right] dt + \alpha_t W_t \sigma_{c,t}^T dZ_t + \alpha_t(\exp(\Delta \ln P_{c,t}) - 1) - C_t dt$$

The lifetime utility of the agent $J(W_t, \tilde{Y}_t)$ is a function of W_t and the state variables for the dynamics, $\tilde{Y}_t$. The agent's HJB equation is:

$$0 = \max_{\{\alpha_t, C_t\}} \min_{P(\eta_t)} f(C_t, J_t) + E_t^{\eta}[dJ]$$

subject to the restriction on $R(\eta_t)$. We are interested in the agent's first-order condition with respect to C_t. Writing out the Lagrangian and taking the derivative with respect to C_t, the FOC is:

$$f_C(C, J) = J_W$$

Homogeneity of the preferences in wealth and linearity of the budget constraint imply that the value function must take the form $J(W, \tilde{Y}) = H(\tilde{Y})\frac{W^\gamma}{\gamma}$ for some function H. Substituting in for $f(C, J)$ and J_W their functional forms, simplifying, and rearranging, one obtains:

$$\frac{C}{W} = H(\tilde{Y})^{\frac{1-\psi}{\gamma}} \delta^\psi \tag{C.1}$$

We want to obtain the consumption-wealth ratio in terms of the function $g(\tilde{Y})$. In equilibrium, the market clears and the agent consumes exactly the aggregate consumption stream, so lifetime utility is given by the equilibrium value of J in (8). Equating the two expressions for J and dividing through by W^γ gives:

$$H(\tilde{Y}) = \exp\left(\gamma g(\tilde{Y})\right)\left(\frac{C}{W}\right)^\gamma$$

Substituting this in for $H(\tilde{Y})$ in (C.1) and solving for $\frac{C}{W}$ gives the result:

$$\frac{C_t}{W_t} = \exp\left(-\rho g(\tilde{Y}_t)\right)\delta \tag{C.2}$$

D 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 (B.1) derived in Appendix B. To facilitate this, let $\hat{A}$ denote the vector obtained by augmenting A with a component equal to 1, 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 (B.1). For the jump terms, we have:

$$\begin{aligned} E_t^\eta\left[\exp\left(G(Y_{t\_} + \xi_t \cdot dN_t)\right) - \exp\left(G(Y_{t\_})\right)\right] =& \frac{\exp\left(\gamma G(Y_t)\right)}{\gamma} E_t^\eta\left[\exp\left(\gamma\hat{A}'\left(\xi_t \cdot dN_t\right)\right) - 1\right] \\ =& \frac{\exp\left(\gamma G(Y_t)\right)}{\gamma} l_t^{\eta\prime}\left(\psi_\eta(\gamma\hat{A}) - 1\right) \end{aligned}$$

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 Appendix A.1, under $P(\eta)$ we have: $\xi_i \sim \Gamma\left(k, \frac{\theta}{1-\theta b}\right)$ and $\xi_k \sim \mathcal{N}\left(\mu + \Delta\mu, \sigma^2 s_\sigma\right)$. Therefore, $\psi_i^\eta(\gamma\hat{A}_i) = \left(1 - \frac{\theta\gamma\hat{A}_i}{1-\theta b}\right)^{-k}$ and $\psi_k^\eta(\gamma\hat{A}_k) = \exp\left(\gamma\hat{A}_k\left(\mu + \Delta\mu\right) + \frac{1}{2}\gamma^2\hat{A}_k^2\sigma^2 s_\sigma\right)$.

Substituting the derived expressions into (B.1) and factoring out the term $\exp(\gamma G(Y_t))$ gives the following:

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

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

We can now proceed with the minimization. Let λ_t be the lagrange multiplier on the constraint $R(\eta_t) \leq \varphi q_t^2$. 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.2. 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:

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

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 make the following conjecture: $\lambda_t = \tilde{\lambda}\exp(\gamma G(Y_t))$ 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 FOC 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 FOC's. 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.2). 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 $g(\tilde{Y}_t)$. To that end, I approximate the term $\exp(-\rho g(\tilde{Y}_t))$ in (D.1) via a log-linearization. As (C.2) shows, $\exp(-\rho g(\tilde{Y}_t))$ is just $\frac{1}{\delta}$ times the equilibrium consumption-wealth ratio. I follow Campbell, Chacko, Rodriguez, and Viciera (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, κ_1 are endogenous to the equilibrium solution of the model. As Campbell,

Chacko, Rodriguez, and Viciera (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 (D.1) becomes $\delta\left(\frac{\kappa_0-1}{\rho}+\kappa_1 A_0\right)+\delta\kappa_1 A'\tilde{Y}_t$.

Given the worst-case model parameters, (D.1) 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 (D.1) (and using the log-linearization), we get the following system:

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

Combining these equations with the FOC's 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.

E Equity Return

I follow the approach of Eraker and Shaliastovich (2008). Let $\ln V_{t+s} = \ln M_{t+s} - \ln M_t + \int_t^{t+s} d\ln R_{m,u}$. The Euler equation implies that V_t is a martingale under the worst-case measure:

$$E_t^\eta[d\ln V_t^c+\frac{1}{2}(d\ln V_t)^2+\exp(\Delta\ln V_t)-1]=0 \quad \text{(E.1)}$$

where $d\ln V_t = d\ln M_t + d\ln R_{m,t}$. Log-linearizing $d\ln R_{m,t}$ around the unconditional mean of $v_{m,t}$ implies (17). Further substituting in the conjecture (18) for $v_{m,t}$ gives (19), which expresses $d\ln R_{m,t}$ in terms of $A_{0,m}$, A_m and the state variables. Substituting the expression for $d\ln R_{m,t}$ into $d\ln V_t$ along with the expression for $d\ln M_t$ (15) gives:

$$d\ln V_t = -\theta\delta dt-(1-\theta)\delta\exp\left(-\rho A_0-\rho A'\tilde{Y}_t\right)dt+\kappa_{0,m}dt-(1-\kappa_{1,m})(A_{0,m}+A_m'Y_t)dt+\chi_m'dY_t$$

where $\chi_m = (-\Lambda+\kappa_{1,m}A_m+\delta_d)$. I now employ the exact same log-linearization (9) to $\exp(-\rho A_0+-\rho A'\tilde{Y}_t)$ and replace it with $\kappa_0+\kappa_1\rho A_0+\kappa_1\rho A'\tilde{Y}_t$. Then substituting $d\ln V_t$

into (E.1) and evaluating the expectation results in the following equation:

$$0 = -\theta\delta dt - (1-\theta)\left[\delta\kappa_0 + \delta\kappa_1\rho A_0\right]dt + \kappa_{0,m}dt - (1-\kappa_{1,m})A_{0,m}dt + \chi_m' E_t^\eta[dY_t^c] + \left[(\theta-1)\delta\kappa_1\rho(\hat{A}-\delta_c) + (\kappa_{1,m}-1)A_m\right]' Y_t dt + \frac{1}{2}\chi_m^T\Sigma_t\Sigma_t^T\chi_m + l_t^{\eta\prime}\left(\psi^\eta(\chi_m)-1\right) \quad \text{(E.2)}$$

We can now use the method of undetermined coefficients. This equation must hold for any value of Y_t, which implies that for each component in Y_t the sum of the terms multiplying it must be 0. Furthermore, the sum of the constant terms must be 0. Thus, the equation implies a system of $n+1$ equations whose solution is the $n \times 1$ vector A_m and the scalar $A_{0,m}$. The solution can be found numerically and verifies the conjectured functional form (18) for $v_{m,t}$.

F Integrated Variance

For convenience, let $* \in \{P, \eta, Q\}$ refer to either the reference, worst-case, or risk-neutral measure, respectively. Equation (19) implies that $E_t^*(d\ln R_{m,t})^2 = B_r'\Sigma_t\Sigma_t' B_r + B_r^{2\prime}\left[E^*(\xi_t^2)\cdot l_t^*\right]$, where B_r^2 denotes the vector obtained by squaring the components of B_r. We want to calculate the expectation of integrated variance: $E_t^*[\int_t^T (d\ln R_{m,s})^2] = \int_t^T E_t^*\left(d\ln R_{m,s}\right)^2$. To that end, it is useful to write $E_t^*(d\ln R_{m,t})^2 = \alpha_0^* + \alpha^{*\prime} Y_t$ where α_0 is a scalar and α is a vector of loadings on the state Y_t. The law of iterated expectations implies that: $\int_t^T E_t^*\left(d\ln R_{m,s}\right)^2 = \int_t^T(\alpha_0^* + \alpha^{*T} E_t^*(Y_t))$ A straightforward expansion of the expression for $E_t^*(d\ln R_{m,t})^2$ shows that:

$$\begin{aligned} \alpha_0^* &= B_r' h B_r \\ \alpha^{*\prime} &= B_r' H B_r + B_r^{2\prime}\mathrm{diag}(E^*(\xi_t^2))l_1^* \end{aligned}$$

where $B_r' H B_r$ denotes a row vector where the i-th component is $B_r' H_i B_r$. From this we see that only α^* differs across the measures.

In order to calculate expectations of future values of Y_t, which is required to calculate the integral, it is easiest to express the dynamics of Y_t in terms of demeaned jump shocks (i.e. using the 'compensated' Poisson processes). The general form of compensated dynamics is:

$$dY_t = \mu^* + \widehat{\mathcal{K}^*} Y_t + \Sigma_t dZ_t^* + \xi_t^* \cdot dN_t^* - E_t^*(\xi_t^* \cdot dN_t^*)$$

where $\widehat{\mathcal{K}^*}$ is the resulting transition matrix which incorporates the uncompensated transition matrix, $\mathcal{K}^*$, and the compensation to the jump terms. A standard calculation then gives that:

$$E_t^*(Y_{t+\Delta t}) = \exp(\widehat{\mathcal{K}^*}\Delta t)Y_t + \widehat{\mathcal{K}^*}^{-1}\left(\exp(\widehat{\mathcal{K}^*}\Delta t) - I\right)\mu^*$$

where I is the identity matrix. A straightforward calculation of the integral in $\int_t^T(\alpha_0^* +$

$\alpha^{*T} E_t^*(Y_t))$ results in the following expression for expected integrated variance:

$$E_t^* \left[\int_t^{t+\Delta t} (d \ln R_{m,s})^2 \right] = \alpha_0^* \Delta t + \alpha^{*\prime} \left[\Theta Y_t + \widehat{\mathcal{K}^*}^{-1} \left[\Theta - I \cdot \Delta t \right] \mu^* \right] \tag{F.1}$$

where $\Theta = \widehat{\mathcal{K}^*}^{-1}(\exp(\widehat{\mathcal{K}}^* \Delta t) - I)$.

Finally, I derive the parameters of the compensated dynamics under the three measures. Recall that $\psi^*(u)$ denotes the stacked vector of moment-generating functions evaluated at the vector u. Then we have that $E_t^*(\xi_t^* \cdot dN_t^*) = \text{diag}(\psi^{*(1)}(0)) l_1^* q_t^2$ where $\psi^{*(1)}(0)$ is the first derivative of $\psi^*(u)$ evaluated at 0. Let δ_q be the selector vector for q_t^2, i.e. $\delta_q' Y_t = q_t^2$. Then denote by $[l_1^*]_q$ a matrix such that $[l_1^*]_q \delta_q = l_1^*$. The transition matrix for the compensated dynamics under P is then given by: $\widehat{\mathcal{K}^P} = \mathcal{K} + \text{diag}(\psi^{(1)}(0))[l_1]_q$ and $\mu^P = \mu$. Under the worst-case model: $\widehat{\mathcal{K}^\eta} = \mathcal{K} + [\Sigma_t h_t / q_t^2]_q + \text{diag}(\psi^{\eta(1)}(0))[l_1^\eta]_q$ and $\mu^\eta = \mu$. The difference from P comes from the drift perturbation and the change in the jump intensity and moment-generating function. Finally, under the risk-neutral measure, $\mathcal{K}^{\mathcal{Q}} = \mathcal{K} + [\Sigma_t h_t / q_t^2]_q - H\Lambda + \text{diag}(\psi^{Q(1)}(0))[l_1^Q]_q$ and $\mu^Q = \mu - h\Lambda$, where $H\Lambda$ denotes an $n \times n$ matrix with k-th column equal to $H_k \Lambda$.

The risk-neutral moment-generating function and jump intensity are determined by the worst-case moment-generating function and jump intensity and the price of risk vector Λ. The moment-generating functions are given by $\psi^Q(u) = \psi^\eta(-\Lambda + u)/\psi^\eta(-\Lambda)$, where the division is componentwise. The jump intensity vector is $l_q^Q = \psi^\eta(-\Lambda) \cdot l_t^\eta$. For a proof see Proposition 5 in Duffie, Pan, and Singleton (2000). The risk-neutral expressions show that in going from η to Q, the change in jump intensities and distributions depends on the prices of risk Λ. Risk-neutralization tilts probabilities towards 'high-price' states of the world. The direction and amount of the 'tilt' depends on the magnitude of Λ. For example, note that if $\Lambda = 0$ the worst-case and risk-neutral quantities are identical.

G Detection Error Probabilities

Detection error probabilities are a useful tool for calibrating model uncertainty that is due to Anderson, Hansen, and Sargent (2003). The detection error probability gives the probability that, using a likelihood-ratio test, a decision maker will incorrectly reject the worst-case model in favor of the reference model based on a data sample of a given length T. This is an important statistic because the agent is exactly worried about the possibility that the data has led him to favor the reference model although the true data-generating process is the worst-case model. I now explain how the detection error probability can be calculated in terms of the Radon-Nikodym process η_t.

The likelihood ratio of the worst-case model to the reference model is exactly given by the Radon-Nikodym derivative η_t. Therefore, the probability at time zero of making a

detection error based on a sample of length T is $\text{Prob}^\eta(\ln \eta_T < 0 | \mathcal{F}_0, \eta_0 = 1)$. Note that the probability is evaluated under the worst-case measure. For illustration, I derive the detection error probability for an i.i.d pure diffusion reference model and then discuss how it can be calculated for the framework in this paper.

As Appendix A.1 shows, in a pure diffusion setting $\ln \eta_T = \int_0^T h_t^T dZ_t - \frac{1}{2}\int_0^T h_t^T h_t dt$. Substituting in $dZ_t = dZ_t^\eta + h_t dt$ gives an expression that is more convenient for evaluation under the worst-case measure:

$$\ln \eta_T = \int_0^T h_t^T dZ_t^\eta + \frac{1}{2}\int_0^T h_t^T h_t dt$$

Now consider the distribution of η_T under the worst-case measure. Taking expectations gives

$$E_0^\eta[\ln \eta_T] = \frac{1}{2}\int_0^T E_0^\eta \left[h_t^T h_t\right] dt = \frac{1}{2}\int_0^T 2\varphi = \varphi T$$

When the reference and worst-case models are i.i.d, h_t is constant. It then follows that $\ln \eta_T$ has a normal distribution with variance $T \times h^T h = 2\varphi T$, i.e. $\ln \eta_T \overset{\eta}{\sim} \mathcal{N}(\varphi T, 2\varphi T)$. The detection error probability is then:

$$\text{Prob}^\eta(\ln \eta_T < 0 | \mathcal{F}_0, \eta_0 = 1) = \text{Prob}\left(\mathcal{N}(0,1) < \frac{-\varphi T}{\sqrt{2\varphi T}}\right) = \text{Prob}\left(\mathcal{N}(0,1) < \frac{-1}{\sqrt{2}}\sqrt{\varphi T}\right)$$

Therefore, in this simple case, the detection error probability is $\Phi(\frac{-1}{\sqrt{2}}\sqrt{\varphi T})$, where Φ is the cdf of the standard normal distribution.

In general, a closed-form expression for the detection error probability is not available since the distribution of $\ln \eta_T$ is not known in closed-form. However, for a general class of specifications that includes the affine setting of this paper, the detection error probability can be calculated numerically via Fourier inversion. As Maenhout (2006) shows, using the expression for η_T from Appendix A.1, one can find the (conditional) characteristic function of η_T in closed-form (up to a system of ODEs). The exact detection error probability can then be calculated numerically via a Fourier inversion. This methodology is similar to the one used to calculate option prices in affine settings, as developed in Duffie, Pan, and Singleton (2000). Maenhout (2006) contains a detailed derivation. In calculating the detection error probabilities for the calibrated model, I set the time-0 value of the state vector equal to its unconditional mean. Except for short samples, the detection-probabilities are relatively insensitive to the time-0 value of the state vector.

References

Anderson, Evan, 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.

Anderson, Evan W., Eric Ghysels, and Jennifer L. Juergens, 2007, The Impact of Risk and Uncertainty on Expected Returns, Working paper, Northern Illinois University, University of North Carolina, and Arizona State University.

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

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein, 2005, Can Standard Preferences Explain the Prices of Out-of-the-Money S&P 500 Put Options, Working paper, University of Minnesota.

Bollerslev, Tim, Michael Gibson, and Hao Zhou, 2008, Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities, Working paper, Duke University and the Federal Reserve Board.

Bollerslev, Tim, and Hao Zhou, 2007, Expected stock returns and Variance Risk Premia, Working paper, Finance and Economics Discussion Series 2007-11, Board of Governors of the Federal Reserve System (U.S.).

Brevik, Frode, 2008, State Uncertainty Aversion and the Term Structure of Interest Rates, Working paper, VU Amsterdam.

Britten-Jones, M., and A. Neuberger, 2000, Option Prices, Implied Price Processes, and Stochastic Volatility, *Journal of Finance* 55(2), 839–866.

Broadie, Mark, Mikhail Chernov, and Michael Johannes, 2007, Model Specification and Risk Premiums: Evidence from Futures Options, *The Journal of Finance* 62, 1453–1490.

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

Campbell, John Y., George Chacko, Jorge Rodriguez, and Luis M. Viciera, 2004, Strategic Asset Allocation in a Continous-Time VAR Model, *Journal of Economic Dynamics and Control* 28.

Carr, Peter, and Dilip Madan, 1999, Option Pricing and the Fast Fourier Transform, *Journal of Computation Finance* Summer 1999.

Carr, Peter, and Liuren Wu, 2007, Variance Risk Premia, *Review of Financial Studies* forthcoming.

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

Demeterfi, K., E. Derman, M. Kamal, and J. Zou, 1999, A Guide to Volatility and Variance Swaps, *Journal of Derivatives* 6, 9–32.

Drechsler, Itamar, and Amir Yaron, 2008, What's Vol Got To Do With It, Working paper, The Wharton School, University of Pennsylvania.

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

Duffie, D., J. Pan, and K. 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, *The Journal of Finance* 56, 1367–1403.

Eraker, Bjorn, 2008, The Volatility Premium, Working paper, University of Wisconsin-Madison.

Eraker, Bjorn, and Ivan Shaliastovich, 2008, An Equilibrium Guide to Designing Affine Pricing Models, *Mathematical Finance* forthcoming.

Foresi, Silverio, and Liuren Wu, 2005, Crash-O-Phobia: A Domestic Fear or Worldwide Concern?, *The Journal of Derivatives* Winter 2005.

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov, 2005, There is a Risk-Return Trade-off After All, *Journal of Financial Economics* 76, 509–548.

Hansen, Lars, and Thomas Sargent, 2007, Fragile Beliefs and the Price of Model Uncertainty, Working paper, .

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

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

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, *The Review of Financial Studies* 18, 131–164.

Maenhout, Pascal J., 2004, Robust Portfolio Rules and Asset Pricing, *The 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.

Oksendal, Bernt, and Agnes Sulem, 2007, *Applied Stochastic Control of Jump Diffusions*. (Springer).

Pan, Jun, 2002, The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study, *The Journal of Financial Economics* 63, 3–50.

Santa-Clara, Pedro, and Shu Yan, 2008, Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options, *Review of Economics and Statistics* forthcoming.

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

Tauchen, George, 2005, Stochastic Volatility in General Equilibrium, Working paper, Duke University.

Trojani, Fabio, and Alessandro Sbuelz, 2008, Asset Prices with Locally-Constrained-Entropy Recursive Multiple-Priors Utility, *Journal of Economic Dynamics and Control* Forthcoming.

Ulrich, Maxim, 2008, Model Uncertainty and Term Premia on Nominal Bonds, Working paper, Goethe University.

Uppal, Raman, and Tan Wang, 2003, Model Misspecification and Underdiversification, *The Journal of Finance* 58, 2465–2486.

Table I
Summary Statistics

	VIX^2	Fut^2	VP
Mean	33.30	22.17	11.27
Median	25.14	14.19	8.92
Std.-Dev.	24.13	22.44	7.61
Maximum	163.4	142.4	59.2
Minimum	9.05	2.66	3.27
Skewness	2.00	2.62	2.39
Kurtosis	8.89	11.10	12.03
AC(1)	0.79	0.65	0.65

Table I presents summary statistics for the integrated variance and variance premium measures. The sample is monthly and covers 1990m1 to 2007m3. VIX^2 is the value of the CBOE's VIX index squared and divided by 12 to convert it into a monthly quantity. Fut^2 is the series of monthly realized variances, measured as the sum over a month of squared 5-minute (log) returns on the S&P 500 futures. VP is the measure of the one-month variance premium, constructed as the difference between VIX^2 and a forecast of next month's realized variance.

Table II
Return Predictability by the Variance Premium

Dependent	Regressors		OLS			Robust Reg.		
	X1	X2	β_1	β_2	$R^2(\%)$	β_1	β_2	$R^2(\%)$
r_{t+1}	VP_t		0.76		1.46	1.12		3.20
	(t-stat)		(2.18)			(2.77)		
r_{t+1}	VP_{t-1}		1.26		4.07	1.21		3.75
	(t-stat)		(3.90)			(2.97)		
r_{t+3}	VP_t		0.86		5.92	0.87		6.09
	(t-stat)		(3.19)			(4.12)		
r_{t+1}	VP_t	$\log(\mathrm{P/E})_t$	1.39	-48.67	8.30	1.81	-50.52	10.77
	(t-stat)		(3.00)	(-3.04)		(4.33)	(-4.36)	
r_{t+1}	VP_{t-1}	$\log(\mathrm{P/E})_t$	2.09	-58.12	13.43	1.98	-57.30	12.61
	(t-stat)		(4.82)	(-3.50)		(4.68)	(-4.85)	

Table II presents return predictability regressions. The sample is monthly and covers 1990m1 to 2007m3. 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. OLS denotes estimates from an ordinary least-squares regression. Robust Reg. denotes estimates from robust regressions utilizing a bisquare weighting function. The reported robust regression R^2s are calculated as the ratio of the variance of the implied regression forecast to the variance of the dependent variable.

Table III

Return Predictability by the Variance Premium: Rolling Regressions

Dependent	Regressors		OLS			Robust Reg.		
	X1	X2	β_1	β_2	$R^2(\%)$	β_1	β_2	$R^2(\%)$
r_{t+1}	VP_t		0.72		2.36	0.95		4.10
	(t-stat)		(3.06)			(2.98)		
r_{t+1}	VP_{t-1}		1.03		4.79	1.03		4.77
	(t-stat)		(3.93)			(3.20)		
r_{t+3}	VP_t		0.71		7.23	0.74		7.80
	(t-stat)		(3.59)			(4.47)		
r_{t+1}	VP_t	$\log(\mathrm{P/E})_t$	1.41	-60.39	12.57	1.49	-54.38	11.72
	(t-stat)		(5.91)	(-4.76)		(4.37)	(-4.37)	
r_{t+1}	VP_{t-1}	$\log(\mathrm{P/E})_t$	1.92	-70.89	18.27	1.82	-66.47	16.35
	(t-stat)		(4.78)	(-4.87)		(5.31)	(-5.28)	

Table III presents return predictability regressions. The sample is monthly and covers 1990m1 to 2007m3. The first 24 data points (months) are used to initialize the rolling regression for the variance forecast, so the effective sample for return prediction is 1992m1 to 2007m3. 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. OLS denotes estimates from an ordinary least-squares regression. Robust Reg. denotes estimates from robust regressions utilizing a bisquare weighting function. The reported robust regression R^2s are calculated as the ratio of the variance of the implied regression forecast to the variance of the dependent variable.

Table IV
Calibration – Model Parameters

Preferences	δ	RRA	ψ	φ	
	$-\ln 0.999$	5	2.0	0.0048	
Δc_{t+1}	$E[\Delta c]$	Φ_c			
	0.0016	0.0066			
x_{t+1}	ρ_x	Φ_x	$l_1(x)$	σ_x	
	-0.025	$0.042 \times \Phi_c$	$1.0/12$	$2.25 \times \Phi_x$	
Δd_{t+1}	$E[\Delta d]$	ϕ	Φ_d		
	0.0016	3	$6.0 \times \Phi_c$		
σ^2_{t+1}	ρ_σ	Φ_σ			
	-0.1	0.30			
q^2_{t+1}	ρ_q	Φ_q	$l_1(q)$	μ_q	ν_q
	-0.2238	0.25	$0.75/12$	1.5	1

Table IV presents the parameters for the reference model used in the model calibration. The calibration results are presented in Table V and Table VII. Parameter values are normalized for a monthly interval, i.e $\Delta t = 1$ is one month.

Table V
Model Calibration Results

Statistic	Data		Model		
			5%	50%	95%
Cashflow Dynamics					
$E[\Delta c]$	1.88	(0.32)	0.91	1.92	2.87
$\sigma(\Delta c)$	2.21	(0.52)	2.02	2.45	3.00
$AC1(\Delta c)$	0.43	(0.12)	0.26	0.46	0.63
$E[\Delta d]$	1.54	(1.53)	-2.01	1.74	5.59
$\sigma(\Delta d)$	13.69	(1.91)	10.35	12.38	14.63
$AC1(\Delta d)$	0.14	(0.14)	0.11	0.31	0.49
$corr(\Delta c, \Delta d)$	0.59	(0.11)	-0.01	0.23	0.45
Returns					
$E[r_m - r_f]$	5.41	(2.09)	3.01	6.13	9.51
$E[r_f]$	0.82	(0.35)	1.06	1.51	1.96
$\sigma(r_m - r_f)$	19.48	(2.35)	15.61	18.09	21.07
$\sigma(r_f)$	1.89	(0.17)	0.65	0.89	1.22
$E[p - d]$	3.15	(0.07)	2.75	2.83	2.90
$\sigma(p - d)$	0.31	(0.02)	0.13	0.16	0.21
$skew(r_m - r_f)$ (M)	-0.43	(0.54)	-0.85	-0.26	0.11
$kurt(r_m - r_f)$ (M)	9.93	(1.26)	3.95	5.68	11.00
Variance Premium					
$\sigma(\text{var}_t^P(r_m))$	17.18	(2.21)	8.56	15.44	29.57
$\sigma(\text{var}_t^Q(r_m))$	24.07	(3.15)	11.14	22.36	45.11
$AC1(\text{var}_t^P(r_m))$	0.81	(0.04)	0.75	0.85	0.92
$AC1(\text{var}_t^Q(r_m))$	0.79	(0.05)	0.74	0.84	0.92
$E[VP]$	11.27	(0.93)	5.67	8.28	13.20
$\sigma(VP)$	7.61	(1.08)	3.25	7.82	15.96
$skew(VP)$	2.39	(0.59)	1.38	2.58	4.22
$kurt(VP)$	12.03	(3.30)	5.06	11.05	25.00
$\beta(1)$	0.76	(0.35)	-0.10	1.13	2.82
$R^2(1)$	1.46	(1.52)	0.03	1.83	8.22
$\beta(3)$	0.86	(0.27)	-0.19	0.94	2.37
$R^2(3)$	5.92	(4.67)	0.05	4.04	18.72
$\beta(6)$	0.49	(0.24)	-0.19	0.78	1.91
$R^2(6)$	3.97	(4.74)	0.04	5.37	27.06

Table V presents (a) consumption and dividend dynamics (b) equity premium and risk-free rate moments (c) moments pertaining to the variance premium. For each statistic the table reports its data and model corresponding values. The data for consumption, dividends, the market return, risk free rate, and price-dividend ratio correspond to the period from 1930 to 2006. The data pertaining to the variance premium is based on monthly data from 1990.1-2007.3. For the model I report finite sample statistics based on 1000 simulations each with the corresponding sample size the same as its data counterpart. For the annual data the statistics are based on time-averaged data. The parameters for calibrating the model are given in Table IV. Standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags.

Table VI

Calibration Results: Simulated HAC Standard Errors

Statistic	**Data**		**HAC 50%**
Variance			
$\sigma(\text{var}_t^P(r_m))$	17.18	(2.21)	**2.23**
$\sigma(\text{var}_t^Q(r_m))$	24.07	(3.15)	**3.69**
$E[VP]$	11.27	(0.93)	**1.04**
$\sigma(VP)$	7.61	(1.08)	**1.44**
$skew(VP)$	2.39	(0.59)	**1.60**
$kurt(VP)$	12.03	(3.30)	**7.22**

Table VI presents the median HAC standard error corresponding to the given subset of Monte-Carlo (model-based) moments from Table V. The sample of HAC standard errors is calculated from the same samples used to generate the model-based statistics in Table V. As for the data-based standard errors, the model-based HAC standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags.

Table VII

Model Calibration Results: Physical and Risk Neutral Variance

Statistic	**Data**		**Model**		
			5%	50%	95%
Integrated Variance $\mathbb{P}$					
$\sigma(\text{var}_t^{\text{P}}(r_m))$	17.18	(2.21)	8.56	15.44	29.57
$AC1(\text{var}_t^{\text{P}}(r_m))$	0.81	(0.04)	0.75	0.85	0.92
$R^2(1)$	0.72	(1.28)	0.01	1.19	6.58
$R^2(3)$	1.87	(3.49)	0.03	2.80	15.62
$skew(\text{var}_t^{\text{P}}(r_m))$	1.90	(0.38)	0.49	1.68	3.22
$kurt(\text{var}_t^{\text{P}}(r_m))$	7.61	(1.60)	3.06	6.72	16.57
$corr(\text{var}_t^{\text{P}}(r_m), VP)$	0.86	(0.05)	0.57	0.84	0.95
Integrated Variance $\mathbb{Q}$					
$\sigma(\text{var}_t^{\text{Q}}(r_m))$	24.07	(3.15)	11.14	22.36	45.11
$AC1(\text{var}_t^{\text{Q}}(r_m))$	0.79	(0.05)	0.74	0.84	0.92
$R^2(1)$	0.98	(1.40)	0.01	1.49	7.18
$R^2(3)$	3.05	(4.21)	0.06	3.41	17.45
$skew(\text{var}_t^{\text{Q}}(r_m))$	2.00	(0.49)	0.85	2.11	3.71
$kurt(\text{var}_t^{\text{Q}}(r_m))$	8.89	(2.26)	3.78	8.49	20.03
$corr(\text{var}_t^{\text{Q}}(r_m), VP)$	0.93	(0.03)	0.76	0.93	0.98

Table VII presents moments pertaining to physical expectations of integrated variance, $\text{var}_t^P(r_m)$, and risk-neutral expectations, $\text{var}_t^Q(r_m)$. Physical-measure expectations represent conditional variance. The risk-neutral expectations represent the VIX. For each statistic the table reports its data and model corresponding values. The data is monthly 1990.1-2007.3. For the model, I report finite sample statistics based on 1000 simulations each with the same sample size as the data. The parameters for calibrating the model are given in Table IV. Standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags.

Table VIII
Comparative Statics Results

Statistic	**Data**		**Model 1-A**			**Model 1-B**		
			5%	50%	95%	5%	50%	95%
Returns								
$E[r_m - r_f]$	5.41	(2.09)	0.61	3.86	7.26	-2.81	0.59	4.10
$E[r_f]$	0.82	(0.35)	1.23	1.68	2.12	1.29	1.73	2.17
$\sigma(r_m - r_f)$	19.48	(2.35)	15.56	17.86	20.55	15.75	18.01	20.63
$\sigma(r_f)$	1.89	(0.17)	0.61	0.82	1.13	0.60	0.81	1.11
$skew(r_m - r_f)$ (M)	-0.43	(0.54)	-0.40	-0.06	0.25	-0.25	0.04	0.34
$kurt(r_m - r_f)$ (M)	9.93	(1.26)	3.66	4.62	7.00	3.54	4.26	5.88
Variance Premium								
$\sigma(\mathrm{var_t}(r_m))$	17.18	(2.21)	8.09	13.90	26.04	7.89	13.33	24.88
$\sigma(\mathrm{var_t^Q}(r_m))$	24.07	(3.15)	8.30	14.57	27.54	7.91	13.42	25.00
$E[VP]$	11.27	(0.93)	0.58	0.84	1.34	0.07	0.10	0.15
$\sigma(VP)$	7.61	(1.08)	0.33	0.80	1.63	0.04	0.09	0.18
$\beta(1)$	0.76	(0.35)	-4.90	7.69	23.47	-95.11	21.45	156.20
$R^2(1)$	1.46	(1.52)	0.01	1.02	6.80	0.00	0.47	4.51
$\beta(3)$	0.86	(0.27)	-5.21	6.44	19.57	-92.31	19.57	135.62
$R^2(3)$	5.92	(4.67)	0.02	2.35	15.54	0.01	1.17	10.47

Table VIII presents a comparative statics exercise for the model given in 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.

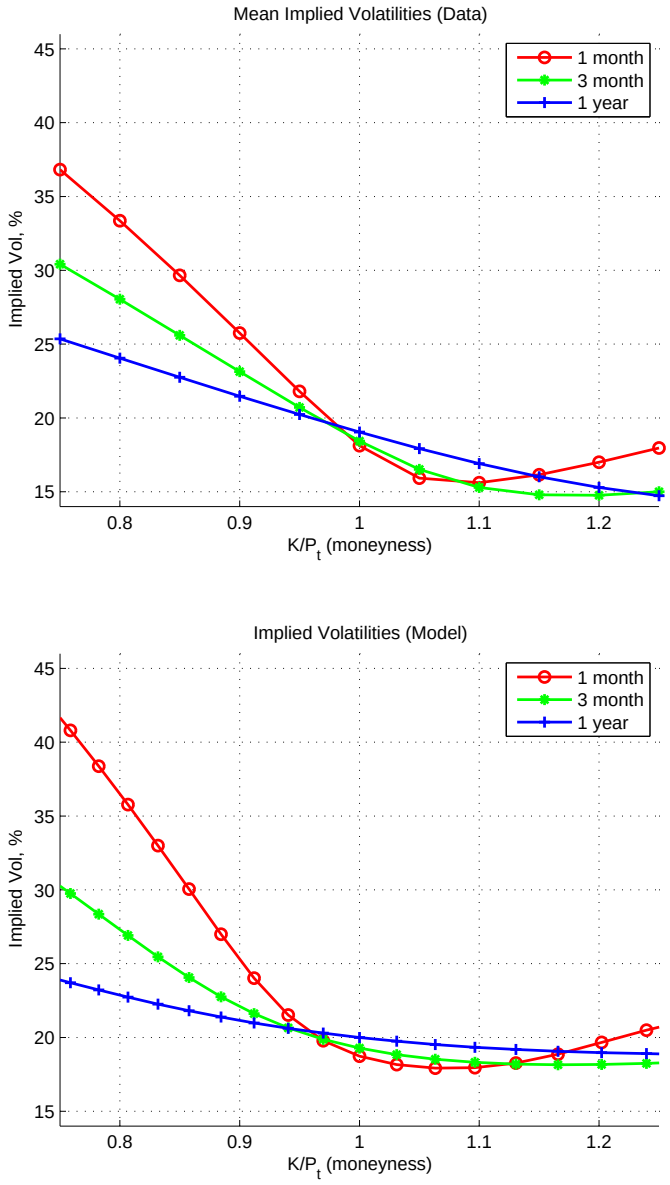
Table IX

An Illustration of the Impact of Time-Varying Uncertainty

Statistic	**Varying Uncertainty**			**Constant Uncertainty**		
	5%	50%	95%	5%	50%	95%
Returns						
$E[r_m - r_f]$	1.00	4.26	7.32	1.59	4.46	7.48
$E[r_f]$	1.31	1.69	2.05	1.17	1.54	1.88
$\sigma(r_m - r_f)$	14.80	16.96	19.70	14.01	15.93	18.10
$\sigma(r_f)$	0.54	0.70	0.89	0.54	0.69	0.89
Variance						
$\sigma(var_t^P(r_m))$	7.69	11.77	20.60	6.53	9.31	13.45
$\sigma(var_t^Q(r_m))$	8.03	13.30	26.41	6.53	9.31	14.45
$E[VP]$	1.01	1.72	3.72	0.01	0.01	0.02
$\sigma(VP)$	0.74	2.36	6.54	0.01	0.01	0.01
$R^2(1)$	0.04	1.62	5.77	0.00	0.38	2.95
$R^2(3)$	0.05	3.52	14.29	0.01	1.03	7.77
$\sigma(\mathrm{var}_t^P(r_m))$ $R^2(1)$	0.00	0.63	4.61	0.00	0.38	2.95
$\sigma(\mathrm{var}_t^Q(r_m))$ $R^2(1)$	0.01	0.82	5.01	0.00	0.38	2.95

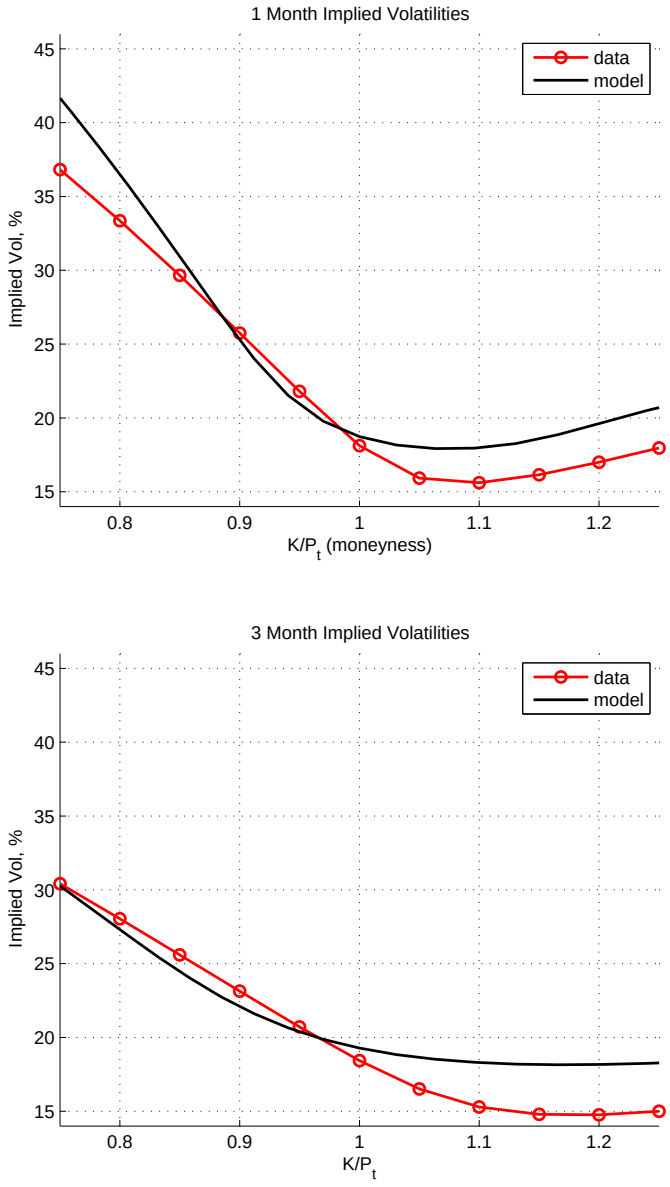
Table IX presents simulation results for the illustration model of Section 6.4.1, which focuses on the impact of time-variation in uncertainty. The left-hand panel shows simulation results for the model where uncertainty is time-varying. The right-hand panel fixes the level of uncertainty at its unconditional mean and raises risk aversion to get the same equity premium. The results are based on 1000 simulations. The simulation parameters are the same as those in Table IV except for the following: $l_1(q) = 1.0/12$, $\mu_q = 2.2$, $\rho_q = -0.313$, $l_1(x) = 0$, $\varphi = 0.0056$. For the right-hand simulations the level of uncertainty is set constant: $\Phi_q = 0$, $l_1(q) = 0$, and risk aversion is 9.

Figure 1: Implied Volatilities: Model and Data



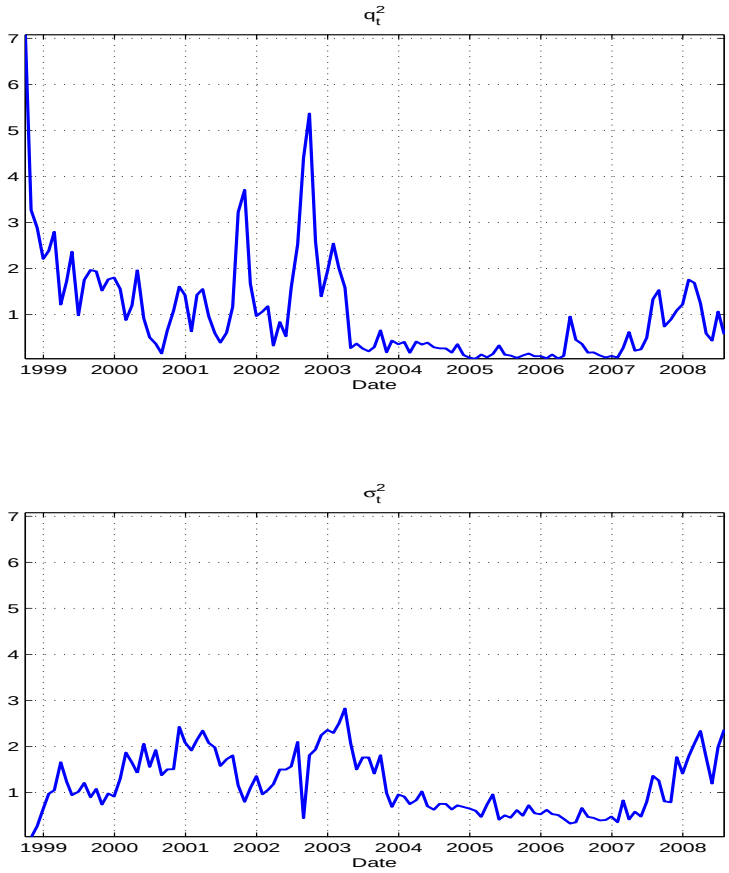
The figure plots implied volatilities from empirical option prices and for option prices calculated for the model of Table IV. The plots show implied-volatility curves for maturities of 1, 3, and 12 months. Strikes are expressed in moneyness (Strike Price/Spot price). The top plot shows the mean of daily implied volatilities for S&P 500 index options for the period 1999.10-2008.6, as quoted in the over-the-counter market. The bottom plot shows the model-based implied volatilities for option prices obtained when the model's state vector is set equal to its unconditional mean.

Figure 2: 1 and 3 month Implied-Volatilities: Model and Data



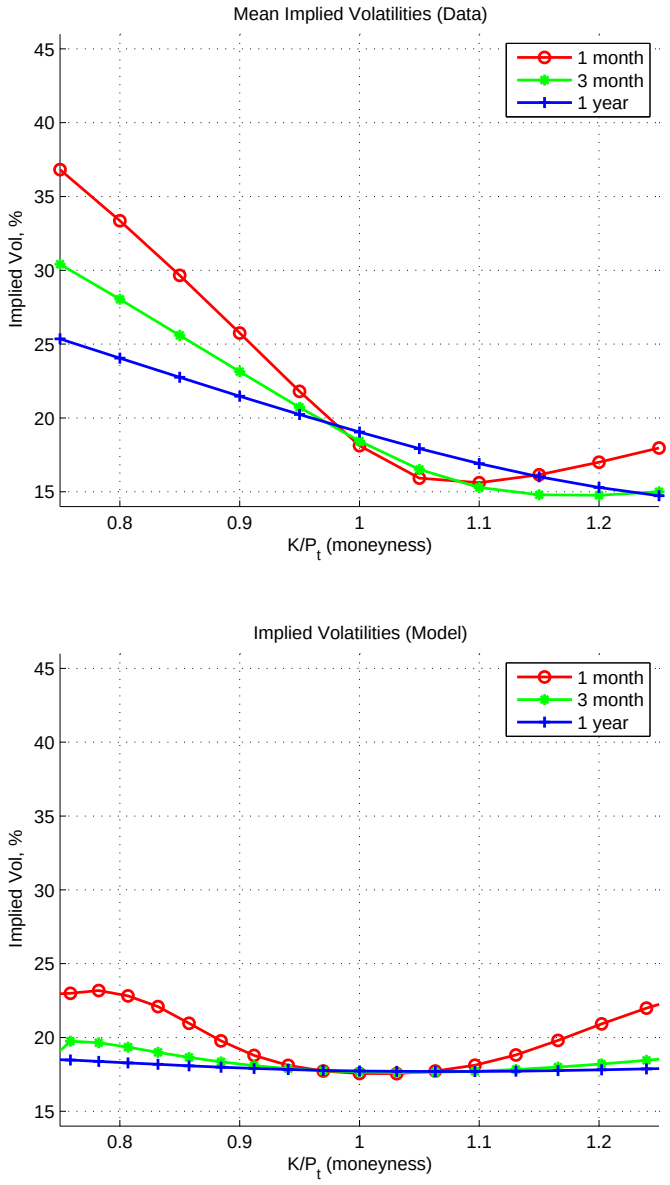
The figure plots comparisons of empirical and model-based implied volatilities for 1 and 3 month maturities for the model of Table IV. Strikes are expressed in moneyness (Strike Price/Spot price). The empirical curves are means of daily implied volatilities for S&P 500 index options for the period 1999.10-2008.6, quoted in the over-the-counter market. The model-based curves are calculated for option prices obtained when the model's state vector is set equal to its unconditional mean.

Figure 3: Option-Implied Time Series of State Variables



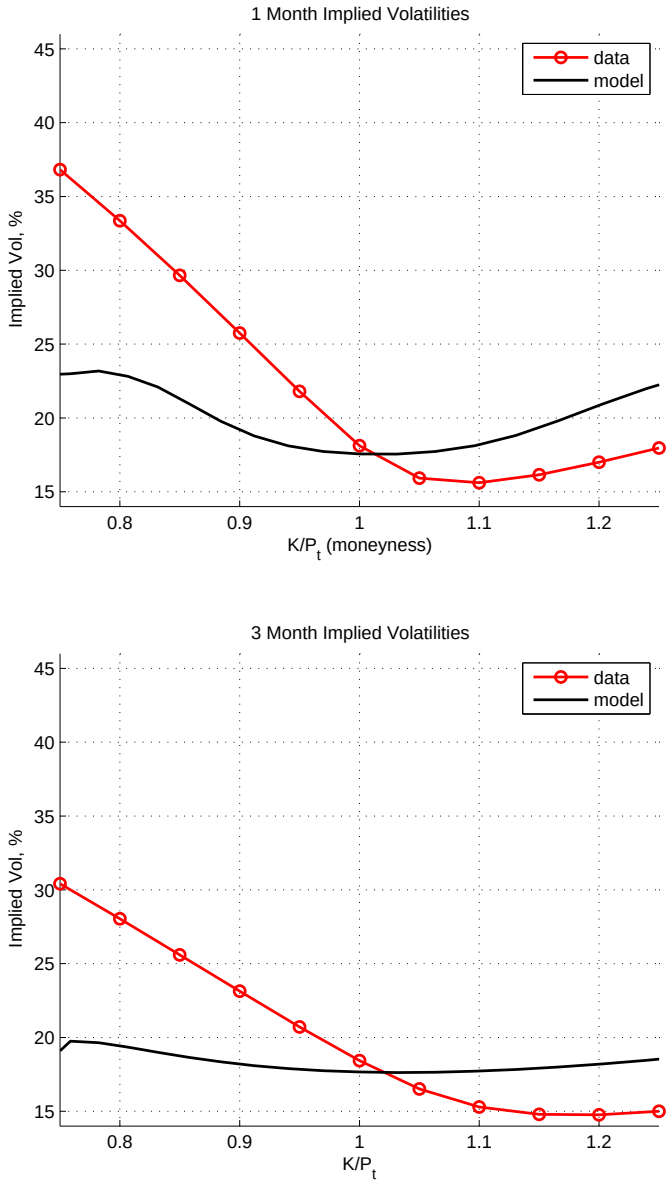
The figure plots the time-series of q_t^2 and σ_t^2 extracted from empirical option prices using the model of Table IV. The sample is 1998.9-2008.7.

Figure 4: Implied Volatilities: No-Uncertainty Model and Data



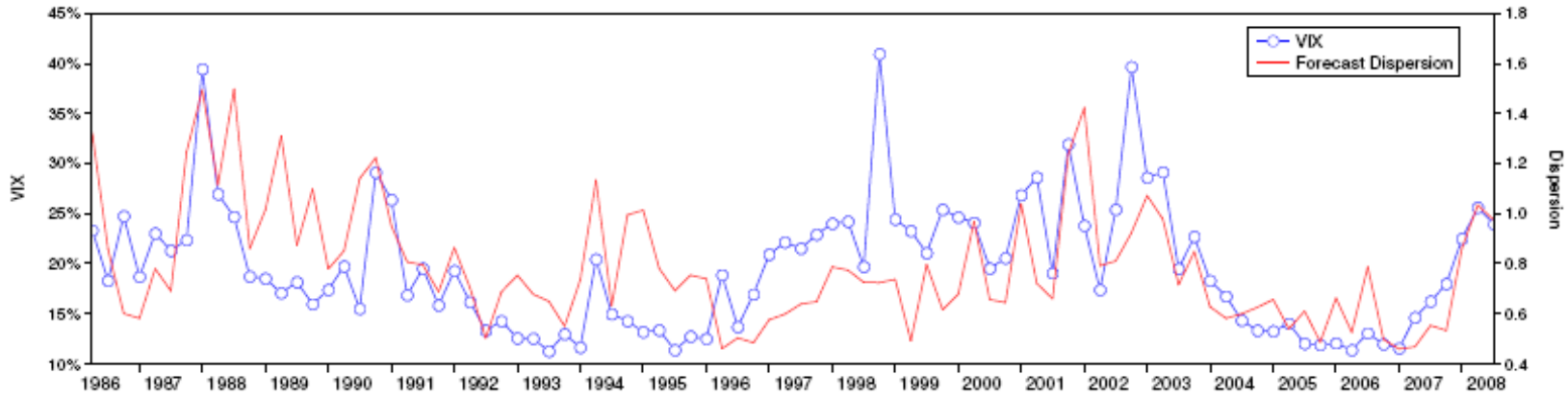
The figure plots the implied volatilities from empirical option prices and for option prices calculated for Model 1-B, which was used in the comparative statics exercise in Table VIII. The plot shows curves for maturities of 1, 3, and 12 months. Strikes are expressed in moneyness (Strike Price/Spot price). The top plot shows the mean of daily implied volatilities for S&P 500 index options for the period 1999.10-2008.6, quoted in the over-the-counter market. The bottom plot shows the model-based implied volatilities for option prices obtained when the model's state vector is set equal to its unconditional mean.

Figure 5: 1 and 3 month Implied-Volatilities: No-Uncertainty Model and Data



The figure plots comparisons of empirical and model-based implied volatilities for 1 and 3 month maturities for Model 1-B used for the comparative statics exercise in 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 1999.10-2008.6, quoted in the over-the-counter market. The model-based curves are calculated for option prices obtained when the model's state vector is set equal to its unconditional mean.

Figure 6: Forecast Dispersion vs. VIX



The figure plots the standard deviation in forecasts of next quarter's real GDP growth from the Survey of Professional Forecasters (SPF) versus the value of the VIX at the end of the previous quarter. The sample is 1986Q2:2008Q3. For the period 1986Q2:1989Q4 the VIX series is actually the VXO (the old measure of the VIX). The correlation of the two series is 0.48 with a standard error of 0.11.