Uncertainty, Time-Varying Fear, and Asset Prices[*]

Itamar Drechsler[†]

First Draft: June 2008
Current Draft: October 24, 2008

Abstract

This paper studies the equilibrium asset pricing implications of time-varying (Knightian) uncertainty regarding economic fundamentals. The paper argues that uncertainty and its variation are important for jointly explaining the equity premium, risk-free rate, and the large variance premium embedded in the "high" price of options. A calibration of the model is able to simultaneously match salient moments of consumption and dividends, the equity premium, risk-free rate, the variance premium and implied-volatility skew, and the documented predictive power of the variance premium for stock returns. The calibration quantitatively demonstrates that uncertainty is strongly reflected in option prices, that fluctuations in the VIX and implied-volatility curve contain an important uncertainty component, and that this component can account for the variance-premium's predictive power. The paper contributes to the ambiguity aversion/robustness literature by solving in closed form for asset prices when the representative agent is ambiguous about both jump and diffusive shocks and has recursive preferences.

Preliminary and Incomplete

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

1 Introduction

Uncertainty or ambiguity about the probabilistic structure of the environment is a different concept than risk, as first pointed out by Knight (1921). Ellsberg (1961) demonstrated that decision-makers have an aversion to uncertainty (ambiguity aversion) that is separate from their aversion to pure risk. Hence, the level of Knightian uncertainty in the economy could have a major influence on asset prices, both unconditionally and through its variation over time. There is evidence that uncertainty does vary over time and that spikes in uncertainty are associated with crises episodes. For example, during 2007-8 forecasters and market participants have repeatedly noted that there is a high, persistent level of uncertainty regarding the real economy.[1] In addition, it has been observed that there is a strong correlation between the level of uncertainty and the VIX index.[2]

This paper builds time-varying uncertainty into a general equilibrium model to quantitatively determine its impact on 'standard' and option-related asset pricing moments. The option-related moments include the variance premium, which is derived from the VIX, and the implied-volatility curve at maturities of one month to a year. A calibration of the model is able to simultaneously match salient moments of consumption and dividends, the risk-free rate, the market return and its volatility, and the option-related quantities. The calibration demonstrates that time-varying uncertainty is particularly important for matching the large variance premium embedded in option prices and the steepness of the implied-volatility curve. The calibration is able to capture the predictive power of the variance premium for equity returns, and consistent with the data, it predicts that option-implied measures, such as the variance premium and VIX, should be superior predictors of excess stock returns than physical (true) expectations of volatility.

---

[1]The following is from an October 6, 2008 article on Reuters, "Charles Evans, president of the Chicago Federal Reserve, said risk evaluations reflect 'substantial uncertainty in the outlook for both growth and inflation.' " and "Evans said real economy activity in the United States would stay sluggish into the new year and that the level of uncertainty about the timing of a pickup in growth, which will depend on improvements in the financial and credit markets, 'is very high.' " (Fed's Evans: Weak Growth To Linger, Inflation Too High) Also, from an October 8, 2008 article on CNBC, "My view is that the volatility and uncertainty are far from over and will persist well into Q4," – Global FX Strategist at BMO Capital Markets (CNBC Guest Blog)

[2]An article from Reuters on October 6, 2008 comments on the VIX reaching a new high, " 'This is absolutely amazing. The elevated VIX is reflecting that people are unsure about every financial relationship they have ever known not only in the U.S. but worldwide.' " (VIX Surges to All-Time High as Credit Fears Spread)

Attempting to address moments of cash flows, equity, and options within a single model is challenging. A main challenge for any equilibrium model that addresses this data is that cross-asset relationships make it difficult to account for the large, volatile (hedging) premium in options without implying an unrealistically high equity premium and return volatility or excessively volatile cash-flow processes. In the model of this paper, time-varying uncertainty is important in accounting for the large, volatile option premium. During periods of high uncertainty, fears of jump shocks to important economic fundamentals are amplified. Options serve as a hedge to both these shocks, and to increases in the level of uncertainty. Hence, their prices incorporate a large hedging premium. The model achieves these results with a reasonable level of uncertainty and a relative risk aversion of only 5.

Uncertainty here takes the the form of model uncertainty or model specification concerns. There is a representative agent who 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 how large the alternative set of models he worries about is, at a given time; when uncertainty increases the set of alternative models expands. [3]

The reference model that the agent considers in this paper is flexible, and includes a persistent component in cash-flow growth rates (long-run risk), moderate jump shocks, and stochastic volatility. This flexibility serves two purposes. First, it allows the model to be realistic enough for the calibration to match a large set of moments of cash flows and prices. Second, the rich structure means model uncertainty is allowed to operate through multiple channels and lets the model solution endogenously reveal how much specification concerns there are about the different parts of the economic dynamics. The amount of concern is determined through a tradeoff between the damage a 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. The calibration shows that infrequent jump shocks

---

[3]This framework and motivation correspond 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 fall into both the Robust Control framework and the Recursive Multiple-Priors Utility of Epstein and Schneider (2003), see e.g. Hansen, Sargent, Turmuhambetova, and Williams (2006).

to important state variables, such as 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 worrisome 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. This 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)). This paper shares that conclusion. A big difference is that the model here is preference-based and derives prices starting from macroeconomic fundamentals. 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 robustness to explaining option prices, this paper is closest to Liu, Pan, and Wang (2005) (LPW), who use uncertainty towards rare events to explain the smirk pattern in index options. There are a number of significant differences between LPW and this paper. The environment in LPW is i.i.d, so it cannot address the conditional moments considered here, such as the return predictability and volatility 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 here is more general, allowing for multiple state variables, uncertainty

in both diffusions and returns, 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, and 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 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 and stochastic uncertainty. This makes return volatility, the VIX, and variance premium be imperfectly correlated and thus 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 $E_t^Q \left[\int_t^{t+1} (d \ln R_{m,s})^2 \right]$ - $E_t \left[\int_t^{t+1} (d \ln R_{m,s})^2 \right]$ where Q denotes the risk-neutral measure and $\ln R_{m,s}$ is the (log) return on the market.

---

[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).

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 can be computed by calculating 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 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. This measure is created by summing the squared five-minute log returns on the S&P 500 futures over a 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 squared VIX and lagged realized variance and construct forecasted series 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 conditional forecasts from the projection, gives the series of one-month variance premium estimates. The projection specification used 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. The model calibration also presents a comparison of the empirical and model-based implied-volatility surfaces. Daily data on the volatility surface is obtained from Citigroup and covers October 1999 to June 2008. The model calibration also requires data on consumption and dividends. 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, futures realized variance, and the variance premium measure (VP). Several characteristics are worth noting. First, all three 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. Note also that the mean variance premium is sizeable in comparison with the other two series and is quite volatile.. This is important, since it shows that risk pricing accounts for a large portion of the risk-neutral expectation of variance (the VIX).

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 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 a 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 as much as 12.4%. 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 theory in this paper. Finally, we 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.

3 General Framework

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 economic dynamics. However, the agent does not fully trust that his model is correct. His model uncertainty or specification 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   Reference Model

Let Y_t denote the n-dimensional vector of state variables. The reference model dynamics follow a continuous-time affine jump-diffusion:

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

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 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 struc-

ture. Specifically,

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

where μ and $\mathcal{K}$ are n and $n \times n$-dimensional respectively. It further assumed that the 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. Then, the assumption can 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 state variable in Y_t. This variable will appear repeatedly throughout the model and has the role of governing variation in the level of uncertainty, as discussed below. It is assumed that

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

Finally, let the jump intensity vector take the form $l_t = l_1 q_t^2$ where l_1 is an n-dimensional vector.[5] The partition of Y_t relates to which subset of the model dynamics the agent is uncertain about. As discussed below, uncertainty about the dynamics of a state variable arises from uncertainty about the probability law for either its diffusive or jump shocks. Rather than making the agent ambiguous about all of the state dynamics, the framework is generalized so that the agent is only uncertain about the dynamics of the subset $Y_{2,t}$.[6]

The specification makes q_t drive variation in the size of shocks about which there is

---

[5] 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 or the model calibration, it is omitted to reduce notational complexity.

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

uncertainty. There are several motivations for this specification. First, it is reasonable that the level of uncertainty rises when there is an elevated risk of large shocks to important state variables. It seems quite plausible that the level of uncertainty be related to economic risk. Second, the correlation of these shocks' volatility with uncertainty is consistent with the contention of this paper that variation in uncertainty is related to variation in the VIX. Finally, this specification also facilitates analytical tractability.

3.2   Alternative Models

The alternative models are defined by their probability measures. A requirement for these probability measures is that they put positive probability on the same events as the reference model (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 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 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^{dZ} perturbs dZ_t and η_t^J perturbs the jumps[7].

---

[7]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.

η_t^{dZ} is defined by the SDE:

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

where h_t is an n-dimensional process and $\eta_0^{dZ} = 1$. From Girsanov's theorem we have $Z^\eta = dZ - \int h_t dt$ is a Brownian motion under $P(\eta)$, which implies that $dY_t^c = [\mu(Y_t) + \Sigma(Y_t)h_t] dt + \Sigma(Y_t)dZ_t^\eta$.[8] 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 $P(\eta)$. Below I discuss the resulting dynamics under $P(\eta)$ 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

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

Thus, η^J perturbs the jump intensity by a factor of $\exp(a)$. For the jump size perturbations, 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. 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 $\psi^\eta(u)$.

---

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

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.[9] The determination of this set is now explained.

3.3 The Size of the Alternative Set

As discussed 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 set of alternative models is defined by placing an upper bound on the growth rate of alternative models' relative entropy[10].

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^{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 shows how the set of alternative models is implicitly defined by an upper bound on the relative entropy growth rate.

Since $\eta = \eta^{dz}\eta^J$, the overall relative entropy growth of $P(\eta)$ is the sum of the relative

---

[9]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 guaranteed to be dynamically consistent.

[10]This approach 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. Hansen, Sargent, Turmuhambetova, and Williams (2006) contains a brief discussion of a similar approach to time-varying alternative sets.

entropy growth rates of η^{dz} and η^J. Moreover, the relative entropy of η^J is the sum of the relative 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 of $P(\eta)$, which is denoted by $R(\eta_t)$. As the Appendix A.2 shows, the expression for $R(\eta_t)$ is in terms of $(h_t, a, \Delta u, s_\sigma, b)$.

We now exploit the link between entropy and statistical proximity to define the set of alternative models that concern 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 difficult enough that it warrants concern that this alternative model (and not the reference model) is the true data generating process. This is known as a model detection error. 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 fully 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, which 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. In calibrating the model, the specific value of φ is chosen to imply a particular model detection error probability. This is discussed in detail in the calibration section and in Appendix H. 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.

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 $\psi(C_s, J_s)$. Therefore, for a given probability model, lifetime utility is given recursively by: $J_t = E_t \left[\int_t^\infty \psi(C_s, J_s) ds \right]$. 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:

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

where E^η denotes expectation taken under the probability measure $P(\eta)$.[11][12] 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 $\psi(C, J)$ is standard:

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

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 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:

$$0 = \min_{P(\eta_t)} \psi(C_s, J_s) + E_t^\eta[dJ] \tag{6}$$

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

---

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

[12]This utility specification is an instance of Epstein and Schneider (2003)'s Recursive Multiple Priors utility. $P(\eta)$ is the set of time-0 probability measures formed from the product of the sets of alternative one-step-ahead dynamics

The solution of this equation gives the worst-case perturbation, η^*, which is needed for asset pricing. In order to understand the added complexity here in terms of solution, 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.[13] Moreover, there is no guarantee that the worst-case dynamics associated with η_t^* permit tractable asset pricing.[14] In this paper, solutions for the worst-case model, value function, and asset prices are found that are in closed-form.

4   Solution

Start by expanding the right side of (6) in terms of the perturbation parameters:

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

where J_t^c is the continuous part of J and the second expectation is over the jumps. We can rewrite the first term by applying Ito's lemma and (2):

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

where J_Y is the gradient of J with respect to Y. The Lagrangian corresponding to the minimization in (6) can now be written as:

$$\psi(C_t, J_t)dt + E_t[dJ^c] + h_{2,t}^T \Sigma_{2,t}^T J_{Y_2} dt + E_t^\eta \left[J(Y_{t_-} + \xi_t dN_t) - J(Y_{t_-}) \right] - \lambda_t \left(\varphi q_t^2 - R(\eta_t) \right) \quad (7)$$

where λ_t is the lagrange multiplier on the (time-t) entropy constraint.

Solving for J now 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.

---

[13]Many papers in this 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

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

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 order in the level of consumption. Consequently, I conjecture 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 an approximate analytical solution 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 for 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 of loadings on the $\tilde{Y}_t$. 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, etc...), and $\tilde{\lambda}$, a constant related to the lagrange multiplier λ_t of the entropy constraint, solve the $(n \times 1)$ system of equations

$$0 = \frac{\delta}{\rho}(\kappa_0 + \kappa_1 \rho A_0 - 1) + \hat{A}^T \mu$$

$$0 = \tilde{Y}_t^T A \delta \kappa_1 + \tilde{Y}_t^T \tilde{\mathcal{K}}^T \hat{A} - \frac{1}{\tilde{\lambda}} \left(\hat{A}^T H_q \hat{A} \right)^{\frac{1}{2}} q_t^2 + \frac{1}{2} \gamma \hat{A}^T \Sigma_t \Sigma_t^T \hat{A} + \frac{1}{\gamma} l_1^{\eta'} \left(\psi^\eta(\gamma \hat{A}) - 1 \right) q_t^2 \quad (11)$$

jointly with the equations giving κ_0, κ_1 and 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. The $(n-1)$ system in (11) must be satisfied for any value of $\tilde{Y}_t$, which implies that the terms multiplying a given element in $\tilde{Y}_t$ must sum to 0. This gives $(n-1)$ equations. 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[15]. The Appendix provides details.

Hence, 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.

4.2 Worst-Case Dynamics

As noted in Proposition 1, the solution to the HJB equation involves finding the parameters for the worst-case model. Recall that the perturbation to the drift is $\Sigma_t h_t = [0, \Sigma_{2,t} h_{2,t}]'$. The derivation in Appendix D show that under the worst-case model:

$$\Sigma_{2,t} h_{2,t} = -\frac{1}{\tilde{\lambda}} \Sigma_{2,t} \Sigma_{2,t}^T A_2 = -\frac{1}{\tilde{\lambda}} H_q A_2 q_t^2 \quad (12)$$

where $\tilde{\lambda}$ is a constant that comes out of the equilibrium solution and is closely related to λ_t (the lagrange multiplier). A number of observations can be made. First, the worst-case drift

---

[15]In some specific cases, analytical solutions can be derived in terms of κ_0, κ_1, whose values must still be found numerically. These cases can be useful for finding starting values for numerical solutions to other cases where analytical solutions are not available.

perturbation is just a linear multiple of the vector A_2. The intuition follows from the fact that the A coefficients are the loadings of utility on the state vector. Hence, A determines how sensitive utility is to a (one unit) change in the drift of each variable and therefore which perturbations are most harmful. Second, the size of the perturbation varies over time with q_t^2. This is the case because variation in uncertainty is associated with variation in q_t^2. Hence, when uncertainty is high, the perturbation is large, and vice versa. Finally, $\tilde{\lambda}$ controls the mean size of the perturbation. A large value of $\tilde{\lambda}$ means the perturbation is small on average.

Appendix D also gives the equations determining the worst-case jump perturbation parameters. For both the drift and jumps the solution for the worst-case follows the same principle: 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. For the calibrated model, the calibration section reports the exact allocation of entropy among the perturbations under the worst-case model.

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 also the case for the jumps. Thus, the worst-case model remains affine, which permits tractable asset pricing.

4.3   The Entropy Penalty and Homotheticity

It is interesting to compare the setup of this paper with other approaches, particularly Maenhout (2004), who solves for equilibrium prices in an i.i.d endowment economy with a robust-control agent that has Duffie-Epstein-Zin utility. Maenhout's extension of Hansen and Sargent's canonical robust-control specification is intended to make the representative agent's value function analytically tractable in the case of non-log utility. For log-utility, the value function (and worst-case model) is solvable in closed-form. However, for non-log utility, the preferences are not homothetic in wealth, which Maenhout points out is important for both tractability. To obtain homotheticity, Maenhout scales the entropy penalty parameter in the robust-control problem by a multiple of wealth raised to $(1 - \text{RRA})$. For analytical

convenience, his specific choice of scaling is $(1 - \text{RRA}) \times J$ (the value function).

This penalty-scaling approach is also adopted by other papers, for example, Liu, Pan, and Wang (2005), Uppal and Wang (2003), and Anderson, Ghysels, and Juergens (2007). Although convenient, Maenhout's penalty-scaling led to some concerns. For example, LPW explicity make an effort to argue "that the choice of a normalization factor does not affect, in any qualitative fashion, ... our main result". Therefore, it is interesting to compare Maenhout's approach with the one in this paper, where the value function *is* homogenous in wealth, though homotheticity *is not* explicitly imposed.

The Lagrangian in (7) is essentially the quantity minimized by 'nature' in a robust control problem, with the multiplier λ_t corresponding to the entropy penalty parameter. Appendix D shows that in the framework of this paper we get $\lambda_t = \gamma J(Y_t)\tilde{\lambda}$, where $\tilde{\lambda}$ is a constant. Mapping $\tilde{\lambda}$ to the penalty parameter, we have that the penalty is *endogenously* scaled by the (state-dependent) term $\gamma J(Y_t)$ – the scale factor used by Maenhout and LPW. Thus, imposing the entropy constraint endogenously leads to homotheticity.

The expression for λ_t also shows under what circumstances it is a constant. This is only the case if $\gamma = 0$ (log preferences), otherwise it includes both the stochastic term $\exp(\gamma g(\tilde{Y}_t))$ and the term $\exp(\gamma \ln C_t)$. If $\gamma < 0$ (RRA > 1) then neglecting $\exp(\gamma \ln C_t)$ implies that the penalty parameter becomes larger as $\ln C_t$ increases, which is akin to diminishing the agent's level of uncertainty.

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. The robust-control/uncertainty aversion literature focuses on 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 robustly 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:

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

where $\theta = \frac{\gamma}{\rho}$ and $dR_{c,t} = \frac{dP_{c,t} + C_t}{P_{c,t}}$ 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]:

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

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:

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

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.

---

[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

[17]$d \ln R_{c,t} = d \ln R_{c,t}^c + \Delta \ln R$ where $d \ln R_{c,t}^c$ 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.

5.1.1 The Risk-free Rate

The risk-free rate, $r_{f,t}$, is $E_t^{\eta}[-\frac{dM_t}{M_t}] = -E_t^{\eta}[d\ln M_t^c] - \frac{1}{2}(d\ln M_t^c)^2 - E_t^{\eta}[\exp(\Delta\ln M_t) - 1]$ where $d\ln M_t^c$ is the continuous part and $\Delta\ln M_t$ the jump-related part. 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. In general, the perturbations decrease expected growth and increase expected variation, which increases the precautionary savings motive. Both effects lower the equilibrium risk-free rate.

5.2 Equity

The return on a share in the stock market is now derived. This is an overview, the full details are left to Appendix F. 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.

The share in 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 $d\ln R_{m,t}$ be the instantaneous market return. Following Eraker and Shaliastovich (2008), 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 linearization constants and are given in the Appendix. This log-linearization is similar to the one used earlier for the wealth-consumption ratio.

Now 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 $B_r = (\kappa_{1,m} A_m + \delta_d)$. 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 of equations gives the equilibrium values for $A_{0,m}$ and A_m and verifies the conjecture for $v_{m,t}$. In general this system of equations admits no analytical solution and must be solved for numerically, much as for the A coefficients in the case of the consumption-wealth ratio. The details of the derivation are given in Appendix F.

5.3 The Equity Premium

Given A_m, one can find the equity premium. This is first determined under the worst-case measure and then adjusted to get an expression under the reference measure. The main expressions are highlighted here and the derivation details are left to Appendix F. As usual, the conditional risk 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'. Let $R_{m,t}$ denote the cumulative return through time t on a trading strategy that reinvests all proceeds. The instantaneous market return is $dR_{m,t}/R_{m,t}$. The Euler equation implies that $E_t^\eta [d(M_t R_{m,t})] = 0$.[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:

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

---

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

where as before a superscript 'c' refers to the continuous part and Δ refers to the jump part. 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 terms deal with the jumps. The last term, in particular, accounts for the jump-related covariance between the market return and pricing kernel[19]. Substituting into the previous expression then gives:

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

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:

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

The effect of model uncertainty shows up explicitly via the h_t term and the perturbed jump intensity l_t^{η} and moment-generating function ψ^{η}. For the jump part, the probability change alters $E_t(\Delta M \Delta R_m)$, usually by decreasing it and lowering the covariance of the market return and pricing kernel. In other words, the market return becomes a worse hedge to the intertemporal marginal rate of substitution, which increases its required expected rate of return.

5.4   The VIX and Variance Premium

This section highlights the main reasons that the risk-neutral expectation of return variance (the squared VIX) contains an important uncertainty-related component and why 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 G.

For notational convenience in what follows, let $* \in \{P, \eta, Q\}$ indicate the physical (P), worst-case (η), or risk-neutral (Q) probability measures. Here I take the reference model as

---

[19] The term is related to the identity: $-\text{cov}(\Delta M, \Delta R) = E(\Delta M)E(\Delta R) - E(\Delta M \Delta R)$

the description of the physical measure. Under the measure $*$ the expectation of integrated return variance from time t to T is $E_t^* \left[\int_t^T (d \ln R_{m,s})^2 \right]$, or taking the expectation inside the integral, $\int_t^T E_t^* (d \ln R_{m,s})^2$. I highlight the reasons why 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 $l_t^* = l_1^* \times q_t^2$ and l_1^* is the mean jump intensity under the $*$ measure ($l_1^P = l_1$, $l_1^\eta = \exp(a)l_t$, and l_1^Q is derived in the Appendix). Note that the jump term depends on the specific measure, while the diffusion term does not. Under typical parameterizations (certainly also the calibration in this paper), the magnitude and frequency of jumps is higher under the worst-case model than the reference model, since such perturbations are harmful to utility. Thus, $E^P(\xi_t^2) < E^\eta(\xi_t^2)$ and $l_1^P < l_1^\eta$. This implies that:

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

Thus, over the first instant, expected variance is higher under η than P due to uncertainty about the frequency and magnitude of jumps. Moreover, since the difference is a multiple of q_t^2, it varies directly with the level of uncertainty. Finally, we see that this component, whose size varies with uncertainty, has a greater weight in $E_t^\eta (d \ln R_{m,t})^2$ than in $E_t^P (d \ln R_{m,t})^2$.

A similar result obtains when comparing this quantity under η and Q. Recall that the agent determines asset prices, and hence the risk-neutral measure, 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 $l_1^\eta < l_1^Q$. This further implies the analog to (21) with Q in place of η and η in place of P. Combining these results gives:

$$E_t^Q (d \ln R_{m,t})^2 - E_t^P (d \ln R_{m,t})^2 = B_r^{2'} \left[E^Q(\xi_t^2) \cdot l_1^Q - E^P(\xi_t^2) \cdot l_1^P \right] q_t^2 > 0 \qquad (22)$$

Thus, for $u = t$, the quantity $E_t^* (d \ln R_{m,u})^2$ is highest for Q and lowest for P. The next question is how this quantity evolves as u increases to some later time s. I sketch the main idea, leaving the exact details to the Appendix. Consider first the difference in the "drift"

23

of this quantity between P and η. Part of the difference is due to the drift perturbation (12) under the worst-case model. This drift 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. Moreover, as (12) shows, the size of the drift perturbation is a multiple of q_t^2 and therefore varies with uncertainty. This causes $E_t^* \left(d \ln R_{m,u} \right)^2$ to have a higher drift under η than P and by an amount that varies 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:

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

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

Again, a similar result again obtains for η and Q. Marginal utility is higher when q_t^2 is higher (since it implies higher uncertainty and jump intensity). Hence, the risk-neutral measure tilts probability mass towards states with higher values of q_t^2, which results in a higher drift in q_t^2 under Q than η. This implies that $\int_t^T E_t^Q \left(d \ln R_{m,s} \right)^2 > \int_t^T E_t^\eta \left(d \ln R_{m,s} \right)^2$. Combining this result with the result for η and P implies that:

$$E_t^Q \left[\int_t^T \left(d \ln R_{m,s} \right)^2 \right] > E_t^P \left[\int_t^T \left(d \ln R_{m,s} \right)^2 \right] \tag{23}$$

The one-month variance premium, $vp_{t,t+1}$, is exactly the difference, for $T = 1$, between the two expectations in (23). Thus, $vp_{t,t+1} > 0$. 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 (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. In these cases, $vp_{t,t+1}$ is a good filter for q_t^2 since it filters out most of the influence of these other volatility drivers, which means variation in it mostly reflects variation in q_t^2.

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:

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

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

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 AR(1) process. Hence, σ_t^2 governs the immediate level of *risk* in cash flow growth rates. Moreover, I assume there is no ambiguity about the structure of these immediate cash flow growth rates. using the notation from Section 3.1, I let $Y_{1,t} = (\ln C_t, \ln D_t)'$ and $Y_{2,t} = (\sigma_t^2, x_t, q_t^2)$. I make q_t^2 follow an AR(1) process. Thus, there is uncertainty about 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 $E(q_t^2) = 1$.[20]. 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

---

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

$\xi_t = (0, 0, 0, \xi_{x,t}, \xi_{q,t})'$. 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. As ν_q decreases, the right tail of the distribution becomes thicker and the distribution becomes more asymmetric. 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 and return 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. Finally, letting q_t^2 drive the diffusion volatility for these state variables facilitates tractability if one wishes to have uncertainty with respect to the diffusion.

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 projections of stock returns on these variance-related quantities. Finally, the calibration also compares the model's 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 III 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 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 III 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 III 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 is 1.5. These jumps capture spikes in the level of uncertainty, which is instrumental in capturing the high variance premium (and high price of options).

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, Appendix H derives an expression that directly links

the level of φ to the detection-error probability for the worst-case model. The detection-error probability is the probability that, if the true dgp is the worst-case model, a likelihood ratio test would actually favor the reference model. It is given by the probability that the log-likelihood ratio of the worst-case and reference models is negative. Models with high detection-error probabilities are difficult for the agent to distinguish empirically from the reference model and therefore present a concern. The detection-error probability depends on the span of the data history the agent uses for inference regarding the dgp. This history is limited by both the availability of data records and also by any structural breaks that may have occurred within the sample. I choose φ to correspond to a 10% detection-error probability for the post-war sample or 7% for the long sample (1930:2006)[21]. Note that two indistinguishable models correspond to a maximum 50% (not 100%) detection-error probability. These are non-trivial chances of detection-error, so concerns about incorrectly rejecting the worst-case model seem quite reasonable. Of course, if one believes the useful sample is shorter, then the detection-error is higher. For comparison, the same φ also implies detection-error probabilities of 22% and 43% for sample lengths of 20 years and 1 year respectively.

6.3 Results

Table IV 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 250 simulations, each with 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 'standard' asset pricing moments, such as the mean and volatility of the market and risk free rate. 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 250 simulations. It is important to note that the *reference* model's dynamics are the ones being simulated. These are the

---

[21]This detection-error probability is approximate since I am using the formula (H.3) for the case of constant ambiguity and volatility.

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 90% confidence intervals generated by the model simulations. For completeness, I also provide HAC robust standard errors of the data statistics.

The top panel in Table IV 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 despite the presence of jump shocks, 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 quite a reasonable specification for agents to use as their reference model.

The second panel of Table IV presents some of the model's asset pricing implications. This panel pertains to 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 low relative risk aversion of 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 despite separating the consumption and dividend processes. Moreover, the model here is able to match 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 is able to replicate the negative skewness and high kurtosis observed in the data[22]. The one moment where the model falls somewhat short is in generating the large volatility of the price-dividend ratio.

---

[22]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 bottom panel in Table IV provides a number of statistics pertaining to integrated variance and the variance premium, all at the monthly horizon. The impact of time variation in model uncertainty shows up very strongly here. First, as in the data, the model has a large variance premium so that the risk-neutral expectation of integrated variance is substantially larger on average than its physical measure counterpart. Second, the model's conditional variance premium is also very volatile, as in the data. The table shows that the model almost exactly matches the volatility of expected integrated variance under the two measures, with the risk-neutral expectation of integrated variance varying substantially more than the physical expectation. Moreover, the model's median skewness and kurtosis for the variance premium are right in line with the large values in the data. Finally, the model's median autocorrelation for the P and Q expectations of integrated variance are close to the data and are easily within the 90% confidence interval. This shows that conditional return volatility inside the model is persistent but not extremely so, as is the case in the data.

There are two main ways that uncertainty helps to produce a large and volatile variance premium. The first reason is that, as discussed in Section 5.4, several of the perturbations in the worst-case model cause an increase in variance. For instance, both the level of the jump intensity and the drift in q_t^2 are increased under the worst-case model. The reason is that both perturbations harm the agent's utility. The jump perturbations, in particular, are both harmful and difficult to detect. This means they are allocated a substantial fraction of the 'entropy budget' and figure prominently in the worst-case model. Moreover, the means of the jump size distributions change under the worst-case model in ways that also increase the expectation of integrated variance. For example, the mean x_t jump size, which is 0 under the reference model, becomes negative under the worst-case model. 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. Since shocks to q_t^2 cause variation, this large, negative risk price increases risk-neutral expectations of integrated variance. The 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. Therefore, the agent wants to hedge increases in q_t^2 and is willing to pay a high premium for assets that pay well when uncertainty spikes up (for example options) . Thus, time variation in uncertainty causes q_t^2 shocks to carry a large, negative risk price, which leads to a large variance premium.

An increase in uncertainty and loss of confidence in the reference model increase the distance between the worst-case and reference models. The agent then perceives that growth

prospects are worse (through the perturbation to x_t) and that the risk of jump shocks is higher (through the perturbation to the jump intensity). Both affects decrease prices and thereby increase expected returns. This implies that high levels of uncertainty are associated with high expected excess returns and that the level of uncertainty has predictive power for returns. Since uncertainty is also the main driver of time variation in the variance premium, as discussed in Section 5.4, the variance premium should be a predictor of excess stock returns. The predictive ability of the variance premium in the data was highlighted in the Data section. As the bottom panel of Table IV 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 3.9% and 5.3% inside the model, which, though high, is similar to the 5.9% and 4.0% values in the data. Overall, the results of Table IV indicate that the model can capture quite well the cash flow, asset pricing and variance-related moments in the data.

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 is consistent with this observation. 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 since its variation is driven almost exclusively by q_t^2. Hence, the variance premium should be a better predictor of excess returns than just the conditional variance.

This point can be seen from Table V, which provides a more in-depth look at the properties of expected integrated variance under the P and Q measures. The top panel displays properties of P-measure expected integrated variance, which corresponds closely to conditional one-month variance. The bottom panel looks at the Q-measure expectation, which effectively replicates the 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 IV, one sees 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, at both maturities in the data and the model. This is because the risk-neutral expectations more strongly reflect the level of q_t^2, which is important for pricing. Therefore the Q-measure expectation does a better job of revealing the level of q_t^2 than does the P-measure expectation, though it is not as good at this as the variance premium. This is further reflected in the skewness and kurtosis statistics of both the model and data, which reflect that variation in q_t^2 has the greatest influence on the variance premium and 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, Table V 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 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 summarizes how 'expensive' options are, one may be interested in the whole volatility surface implied by the model. To that end, I price options within the calibrated model calculate their Black-Scholes implied-volatilities. I numerically calculate option prices by using Fourier-transform techniques. I find the characteristic function for the state vector using the methods in Duffie, Pan, and Singleton (2000) and then calculate and invert the option transform based on the method of Madan and Carr

(1999). The details of this procedure are in Appendix I. Figure 1 shows a comparison of the model-based and empirical implied volatility curve 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 wide range of strikes for short-maturity options. The empirical data is obtained is obtained from Citigroup and covers October 1999 to June 2008. It represents the averge of daily implied volatilities on over-the-counter options on the S&P 500 index. The range of strikes available in the otc market for index options is often much wider than for exchange-traded options.See ? 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 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 does fit is very good for low moneyness strikes but not as good for high-moneyness.

6.4   The Impact of Ambiguity

Table VI conducts a two-part comparative statistics exercise on the model of Table III by shutting off ambiguity with respect to parts of the model's dynamics. The first panel, labeled Model 1-A, is for a model that shuts off ambiguity with regards to only the jump shocks in the model, leaving on ambiguity regarding the diffusion part of the dynamics. Thus, the jump components are unchanged under the corresponding 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

different 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 significantly reduces the median equity premium, though the premium remains non-trivial. The other moments in the top panel are not greatly affected. 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 very 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 variance premium is reduced by roughly 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 variance premium even further, so that it is essentially zero. Finally, 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, as 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.

Finally, the results from the table are reinforced by Figures 3 and 4, which plot the implied-volatility curves for Model 1-B. The figures show that the model-implied curves are very flat and the model does not capture the steep skew in implied-volatility. This is particularly at apparent the 1-month maturity, though it is also a problem at maturities of 3 and 12 months.

6.5   Forecast Dispersion and the VIX

Figure 5 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. 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 causes a sharp spike in the VIX without a corresponding strong increase in forecast dispersion. For completeness, I also calculate the model's correlation between the level of q_t^2 and the model-implied VIX. For the calibrated model simulations described above, the 5, 50, and 95 percentiles of this correlation are 52%, 78%, and 94% 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 part of studying asset prices in equilibrium models is that the 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 high mean excess equity returns and the large variance premium embedded in option prices. Model uncertainty or concerns for robustness present an intuitively appealing and promising direction for explaining these high risk premia. However, to explore this avenue effectively requires building models with a structure rich enough to enable robustness 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 presents a flexible, multivariate framework for solving for equilibrium asset prices with uncertainty aversion/concern for model robustness. Uncertainty is modeled as a local constraint on the relative entropy of alternative models and is allowed to vary over time. The paper illustrates how to solve for the worst-case model within the multivariate framework and further shows that asset pricing remains tractable under the resulting worst-case model. As an additional point, the paper illustrates how the framework used here is related the 'homothetic' robustness technique of Maenhout (2004).

The paper applies this framework to introduce time-varying uncertainty into an economy with long-run risks, stochastic volatility, and jump shocks. 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. Uncertainty regarding these dynamics has important asset pricing implications. The paper presents a calibration of the model economy that is able to jointly match moments of cash flows, the equity premium and risk-free rate, and option prices. Uncertainty is particularly important for matching the large and volatile variance premium embedded in option prices and the high implied volatilities of out-of-the-money put options. The calibration further shows that uncertainty regarding jump shocks is important for capturing the return predictability of equity returns by option-implied quantities, such as the variance premium and VIX. Finally, the calibration demonstrates that these 'priced' variance measures reflect uncertainty more strongly than statistical variance measures, which can explain their greater predictive power for returns.

Appendix

A Measure Change and Entropy for Jumps

A.1 Derivation of Measure Changes for Jumps

TBA

A.2 Derivation of Relative Entropy Growth for Jumps

TBA

B HJB Derivation

TBA

C Equilibrium Consumption-Wealth Ratio

In equilibrium, markets clear so that the representative agent must hold all of his wealth in the claim to aggregate consumption. Thus, the equilibrium consumption-wealth ratio can also be viewed as the dividend-price ratio of the aggregate consumption claim.

To derive the equilibrium consumption-wealth ratio, consider the consumption and portfolio problem of the representative agent in this Lucas-tree setting. Since the information filtration is generated by Z_t, we can assume that the price of the aggregate consumption claim P_c, follows an Itô process:

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

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:

$$dW_t = W_t \left[\alpha_t(u_t - r_{f,t}) + r_{f,t}\right]dt + W_t\alpha_t\sigma_{c,t}^T dZ_t - C_t dt$$

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:

$$0 = \min_{h_t} \ \sup_{\{\alpha_t, C_t\}} \ \psi(C_t, J_t) + E_t^h[dJ]$$

subject to the restriction on the norm of h_t. The only part of this problem used at this point, is the agent's first-order condition with respect to C_t. Writing the Lagrangian and taking the derivative with respect to C_t, the FOC is:

$$\psi_C(C, J) = J_W$$

By the usual argument, 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 $\psi(C, J)$ and J_W their functional forms, simplifying, and rearranging, one obtains:

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

To get the consumption-wealth ratio in terms of the known function $g(\tilde{Y})$, instead of $H(\tilde{Y})$, we use the expression for the equilibrium value function (8). In equilibrium, since the market clears and the agent consumes exactly the aggregate consumption stream, lifetime utility equals J given in (8). Equating the two expressions for J and dividing through by W^γ implies:

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

Substituting this in for $H(\tilde{Y})$ in the expression above for $\frac{C}{W}$ and solving for $\frac{C}{W}$ leads to the result:

$$\frac{C_t}{W_t} = \exp\left(-\rho g(\tilde{Y}_t)\right) \delta \tag{C.1}$$

D   Proof of Proposition 1

TBA

E   Numerical Solution Details

TBA

F   Equity Return

TBA

G   Integrated Variance

TBA

H   Detection Error Probability

The detection error probability is the probability that the wrong model/dgp will be inferred from the data history. The model chosen is the one with the higher likelihood, and the comparison is between the worst-case and reference models. The detection error probability is the answer to the question: what is the probability that the true dgp is the worst-case model, but a likelihood ratio has (erroneously) led the agent to believe the reference model is the dgp.

Let $\{Y_t\}_{t \in [0,T]}$ be the history of state vector realizations. The probability of this history is the probability of the history of innovations dZ_t that produced it (conditioning on the initial observation). The Radon-Nikodym derivative η_T is exactly the likelihood ratio for the worst-case model relative to the reference model for this history of innovations. Therefore, the detection error probability is $\text{Prob}^h(\ln \eta_T < 0)$, where Prob^h denotes the probability under the worst-case measure. From $\frac{d\eta_t}{\eta_t} = h_t^T dZ_t$ and $\eta_0 = 1$, Itô's lemma implies that $\eta_T = \exp\left(\int_0^T h_t^T dZ_t - \frac{1}{2}\int_0^T h_t^T h_t dt\right)$ and therefore $\ln \eta_T = \int_0^T h_t^T dZ_t - \frac{1}{2}\int_0^T h_t^T h_t dt$. Substituting in $dZ_t = dZ_t^h + h_t dt$ gives an expression that is more convenient for evaluation under the worst-case measure:

$$\ln \eta_T = \int_0^T h_t^T dZ_t^h + \frac{1}{2}\int_0^T h_t^T h_t dt \tag{H.1}$$

Now consider the distribution (from time 0 viewpoint) of η_T, under the worst-case measure. It is easy to see that

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

If h_t is constant, which is the case for the constant volatility and ambiguity specification discussed above, then it follows that $\ln \eta_t$ has a normal distribution with variance $h^T h T =$

$2\varphi T$, i.e. $\ln \eta_T \overset{h}{\sim} \mathcal{N}(\varphi T, 2\varphi T)$. The detection error probability is

$$\text{Prob}^h (\ln \eta_T < 0) = \text{Prob}\left(\mathcal{N}(0,1) < \frac{-\varphi T}{\sqrt{2\varphi T}}\right) = \text{Prob}\left(\mathcal{N}(0,1) < \frac{-1}{\sqrt{2}}\sqrt{\varphi T}\right) \qquad \text{(H.3)}$$

Therefore, in this 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 the detection error probability can be found in closed-form, up to a (numerical) Fourier inversion. For example, for the stochastic volatility and ambiguity specification discussed above, the expression $\frac{1}{2}h_t^T h_t + h_t^T dZ_t^h$ (which appears under the integral in (H.1)), is affine in σ_t^2 [23]. In this case, Maenhout (2006) shows, using results from Duffie, Pan, and Singleton (2000), that it is possible to solve for the characteristic function of $\ln \eta_T$ and then numerically invert it to get the exact detection error probability. Finally, Monte-Carlo simulation of the model is a less elegant, but even more general method of calculating detection error probabilities.

I Calculating Option Prices by Fourier Transform

TBA

---

[23]Specifically, it can be written as $a_1\sigma_t^2 dt + \sigma_t a_2^T dZ_t$ for some constant coefficients a_1 and a_2

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.

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.

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

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.

Ellsberg, Daniel, 1961, Risk, Ambiguity, and the Savage Axioms, *The Quarterly Journal of Economics* 75, 643–669.

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

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.

Knight, Frank, 1921, *Risk, Uncertainty, and Profit.* (Houghton Mifflin Company).

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.

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

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.

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
AR(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 CBOE's VIX squared and divided by 12 to convert it into a monthly quantity. Fut$^2_{t+1}$ is the sum over a month of squared 5-minute returns on the S&P 500 futures. VP is the measure of the variance premium described in the text.

Table II

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	$\log{(\text{P/E})}_t$	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}	$\log{(\text{P/E})}_t$	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 total 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.

Table III

Calibration – Model Parameter Configuration

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_{t+1}^2	ρ_σ	Φ_σ			
	-0.1	0.30			
q_{t+1}^2	ρ_q	Φ_q	$l_1(q)$	μ_q	ν_q
	-0.2238	0.25	$0.75/12$	1.5	1

Table III presents the parameters for the reference model used in the model calibration.

Table IV
Model Calibration Results

Statistic	Data		Model		
			5%	50%	95%
Cashflow Dynamics					
$E[\Delta c]$	1.88	(0.32)	0.94	1.95	2.82
$\sigma(\Delta c)$	2.21	(0.52)	2.02	2.45	3.00
$AC1(\Delta c)$	0.43	(0.12)	0.27	0.46	0.63
$E[\Delta d]$	1.54	(1.53)	-1.77	1.45	5.31
$\sigma(\Delta d)$	13.69	(1.91)	10.62	12.39	14.65
$AC1(\Delta d)$	0.14	(0.14)	0.10	0.31	0.48
$corr(\Delta c, \Delta d)$	0.59	(0.11)	0.01	0.23	0.44
Returns					
$E[r_m - r_f]$	5.41	(2.09)	3.05	6.00	9.50
$E[r_f]$	0.82	(0.35)	1.07	1.52	1.96
$\sigma(r_m - r_f)$	19.48	(2.35)	15.65	18.08	21.19
$\sigma(r_f)$	1.89	(0.17)	0.63	0.87	1.27
$E[p - d]$	3.15	(0.07)	2.76	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)	-0.86	-0.27	0.11
$kurt(r_m - r_f)$ (M)	9.93	(1.26)	3.95	5.75	11.55
Variance Premium					
$\sigma(\text{var}_t^P(r_m))$	17.18	(2.21)	9.01	15.57	30.82
$\sigma(\text{var}_t^Q(r_m))$	24.07	(3.15)	11.95	22.47	45.94
$AC1(\text{var}_t^P(r_m))$	0.81	(0.04)	0.76	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)	5.81	8.23	13.34
$\sigma(VP)$	7.61	(1.08)	3.47	7.75	17.76
$skew(VP)$	2.39	(0.59)	1.32	2.68	4.25
$kurt(VP)$	12.03	(3.30)	4.82	11.24	25.02
$\beta(1)$	0.76	(0.35)	-0.02	1.13	2.84
$R^2(1)$	1.46	(1.52)	0.03	1.86	8.32
$\beta(3)$	0.86	(0.27)	-0.10	0.88	2.39
$R^2(3)$	5.92	(4.67)	0.04	3.91	18.50
$\beta(6)$	0.49	(0.24)	-0.15	0.75	1.87
$R^2(6)$	3.97	(4.74)	0.04	5.32	27.84

Table IV presents (a) consumption and dividend dynamics (b) asset pricing 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 250 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 III. Standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags.

Table V

Model Calibration Results: Physical and Risk Neutral Variance

| Statistic	Data		Model		
			5%	50%	95%
Integrated Variance $\mathbb{P}$					
$\sigma(\text{var}_\text{t}^\text{P}(r_m))$	17.18	(2.21)	9.01	15.57	30.82
$AC1(\text{var}_\text{t}^\text{P}(r_m))$	0.81	(0.04)	0.76	0.85	0.92
$R^2(1)$	0.72	(1.28)	0.01	1.17	6.40
$R^2(3)$	1.87	(3.49)	0.13	2.59	14.09
$skew(\text{var}_\text{t}^\text{P}(r_m))$	1.90	(0.38)	0.55	1.73	3.32
$kurt(\text{var}_\text{t}^\text{P}(r_m))$	7.61	(1.60)	3.09	6.87	17.19
$corr(\text{var}_\text{t}^\text{P}(r_m), VP)$	0.86	(0.05)	0.57	0.84	0.96
Integrated Variance $\mathbb{Q}$					
$\sigma(\text{var}_\text{t}^\text{Q}(r_m))$	24.07	(3.15)	11.95	22.47	45.94
$AC1(\text{var}_\text{t}^\text{Q}(r_m))$	0.79	(0.05)	0.74	0.84	0.92
$R^2(1)$	0.98	(1.40)	0.01	1.54	7.05
$R^2(3)$	3.05	(4.21)	0.10	3.26	16.38
$skew(\text{var}_\text{t}^\text{Q}(r_m))$	2.00	(0.49)	0.86	2.16	3.57
$kurt(\text{var}_\text{t}^\text{Q}(r_m))$	8.89	(2.26)	3.74	8.70	20.06
$corr(\text{var}_\text{t}^\text{Q}(r_m), VP)$	0.93	(0.03)	0.77	0.93	0.98

Table V presents moments pertaining to physical expectations of integrated variance, $\text{var}_t^P(r_m)$, and risk-neutral expectations, $\text{var}_t^Q(r_m)$. The 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 250 simulations each with the same sample size as the data. The parameters for calibrating the model are given in Table III. Standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags.

Table VI

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)	0.68	3.63	7.22	-2.55	0.50	4.04
$E[r_f]$	0.82	(0.35)	1.25	1.68	2.12	1.30	1.73	2.17
$\sigma(r_m - r_f)$	19.48	(2.35)	15.55	17.90	20.51	15.84	18.13	20.58
$\sigma(r_f)$	1.89	(0.17)	0.60	0.81	1.12	0.59	0.80	1.11
$skew(r_m - r_f)$ (M)	-0.43	(0.54)	-0.39	-0.05	0.22	-0.22	0.05	0.35
$kurt(r_m - r_f)$ (M)	9.93	(1.26)	3.56	4.69	7.71	3.50	4.30	5.84
Variance Premium								
$\sigma(\mathrm{var}_t(r_m))$	17.18	(2.21)	8.55	14.01	27.60	8.44	13.46	26.34
$\sigma(\mathrm{var}_t^Q(r_m))$	24.07	(3.15)	8.81	14.65	28.99	8.45	13.55	26.54
$AC1(\mathrm{var}_t(r_m))$	0.81	(0.04)	0.76	0.86	0.92	0.77	0.86	0.92
$AC1(\mathrm{var}_t^Q(r_m))$	0.79	(0.05)	0.76	0.86	0.92	0.77	0.86	0.92
$E[VP]$	11.27	(0.93)	0.59	0.84	1.36	0.08	0.12	0.19
$\sigma(VP)$	7.61	(1.08)	0.35	0.79	1.81	0.05	0.11	0.26
$\beta(1)$	0.76	(0.35)	-4.82	7.43	23.59	-70.80	18.03	126.8
$R^2(1)$	1.46	(1.52)	0.01	1.04	6.79	0.01	0.45	4.37
$\beta(3)$	0.86	(0.27)	-4.01	5.79	20.15	-61.5	14.60	110.50
$R^2(3)$	5.92	(4.67)	0.02	2.24	14.71	0.01	1.27	10.45

Table VI presents comparative statics exercise for the model given in Table III. The two panels alter the model in TableIII 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.
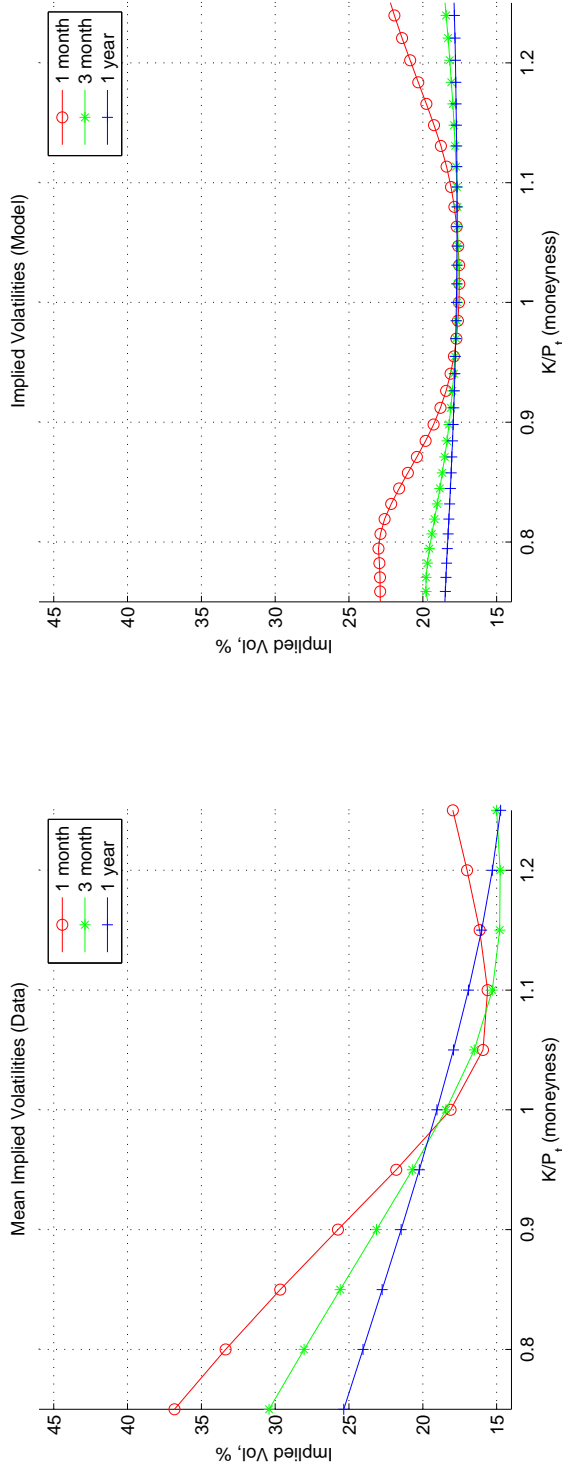
Figure 1: Implied Volatilities: Model and Data
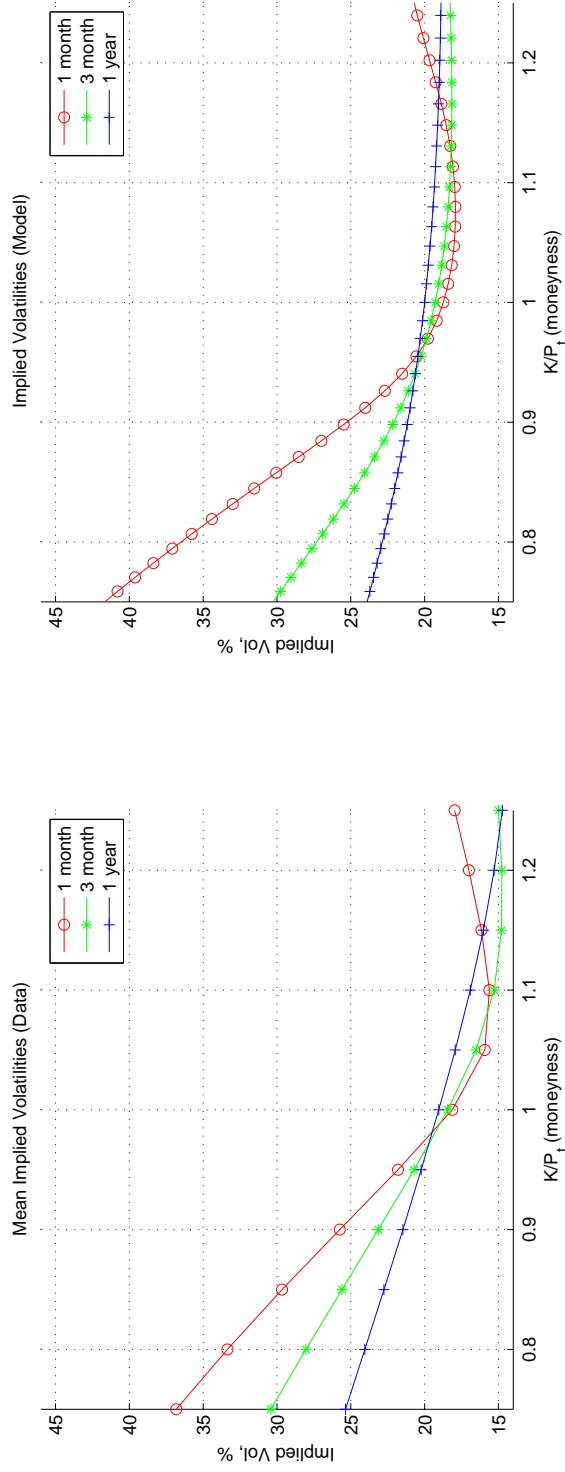
The figure plots implied volatilities from empirical option prices and prices calculated for the model of Table III. The plot shows implied-volatility curves for maturities of 1, 3, and 12 months. Strikes are expressed in moneyness (Strike Price/Spot price). The left plot shows the mean of daily implied volatilities for S&P 500 index options for the period 1999.10-2008.6, quoted in the over-the-counter market. The right plot shows the model-based implied volatilities for option prices that are calculated by setting the model's state vector to its unconditional mean.

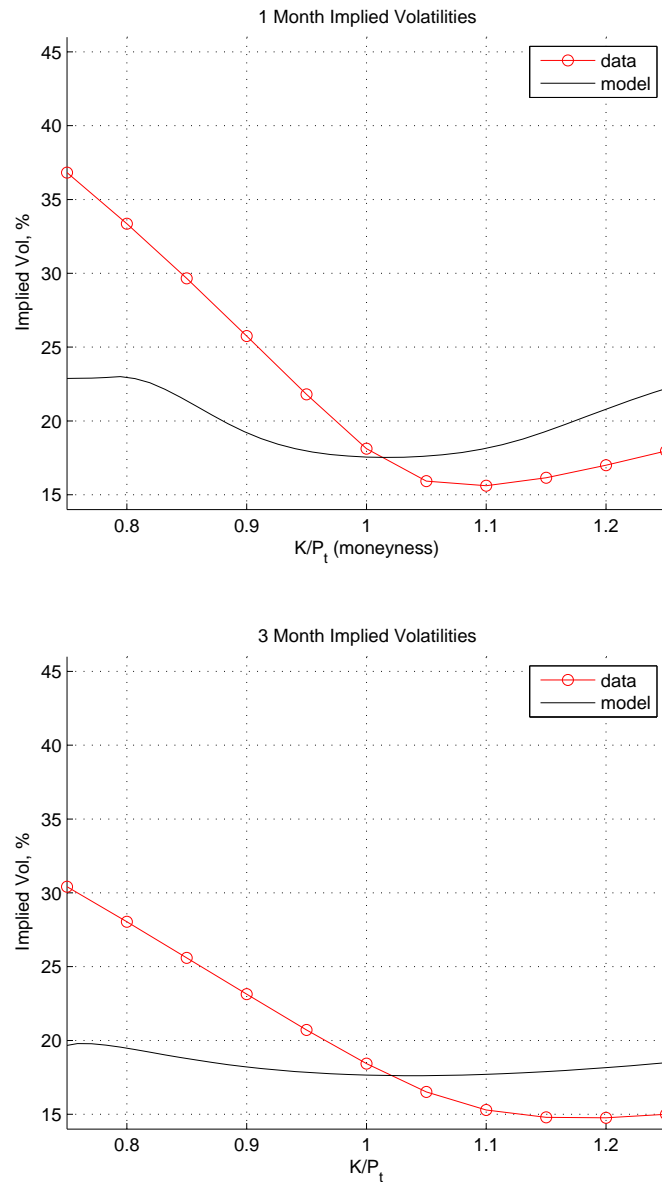Figure 2: 1 and 3 month Implied-Volatilities: Model and Data
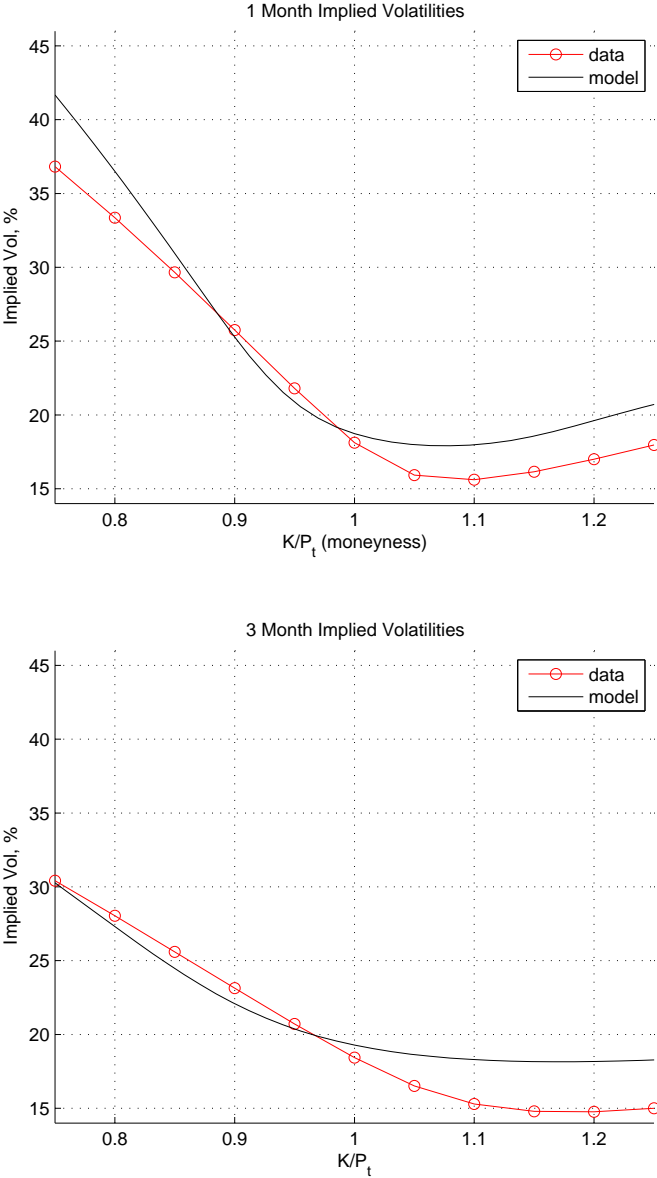
The figure plots comparisons of empirical and model-based implied volatilities for 1 and 3 month maturities for the model of Table III. Strikes are expressed in moneyness (Strike Price/Spot price). The empirical curves are means of daily implied volatilities for S&P 500 index options for the period 1999.10-2008.6, 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 to its unconditional mean.

Figure 3: Implied Volatilities: No–Uncertainty Model and Data

The figure plots the implied volatilities from options both empirically and for the Model 1-B used for the comparative statics exercise in Table VI. The plot shows curves for maturities of 1, 3, and 12 months. Strikes are expressed in moneyness (Strike Price/Spot price). The left 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 right plot shows the model-based implied volatilities for option prices that are calculated by setting the model's state vector to its unconditional mean.

Figure 4: 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 III. 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 to its unconditional mean.

Figure 5: Forecast Dispersion vs. VIX
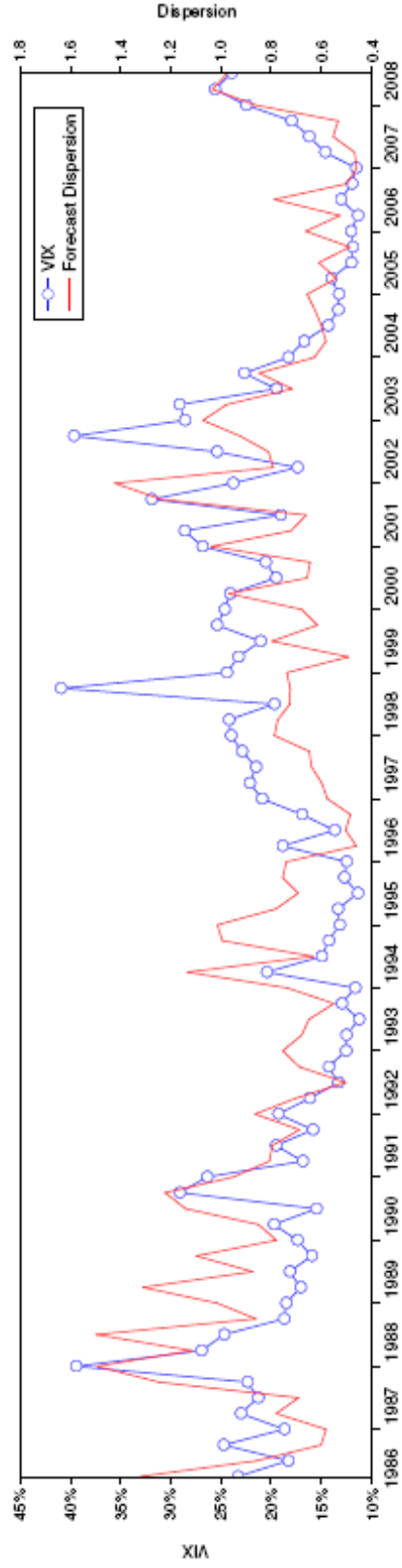
The figure plots the standard deviation in forecasts of next quarter's real GDP growth from the Survey of Professional Forecasters (SPF) versus the value of the VIX at the end of the previous quarter. The sample is 1986Q2:2008Q3. For the period 1986Q2:1989Q4 the VIX series is actually the VXO (the old measure of the VIX). The correlation of the two series is 0.48 with a standard error of 0.11.