What's Vol Got to Do With It*

Itamar Drechsler†

Amir Yaron‡

First Draft: July 2007

Current Draft: November 2007

Abstract

Uncertainty plays a key role in economics, finance, and decision sciences. Financial markets, in particular derivative markets, provide fertile ground for understanding how perceptions of economic uncertainty and cashflow risk manifest themselves in asset prices. We demonstrate that the variance premium, defined as the difference between the squared VIX index and expected realized variance, captures attitudes toward uncertainty. We show conditions under which the variance premium displays significant time variation and return predictability. A calibrated, generalized Long-Run Risks model generates a variance premium with time variation and return predictability that is consistent with the data, while simultaneously matching the levels and volatilities of the market return and risk free rate. Our evidence indicates an important role for transient non-Gaussian shocks to fundamentals that affect agents' views of economic uncertainty and prices.

Preliminary and Incomplete

*We thank participants at the Wharton Lunch Seminar. Drechsler gratefully acknowledges the financial support of the Rodney White Center at the Wharton School.

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

‡The Wharton School, University of Pennsylvania and NBER, yaron@wharton.upenn.edu.

1 Introduction

That idea that volatility has a role in determining asset valuations has long been a cornerstone of finance. Volatility measures, broadly defined, are considered to be a useful tool for capturing how perceptions of fundamental economic uncertainty manifest themselves in prices. The derivatives markets, where volatility plays a prominent role, are therefore especially relevant for unraveling the connections between uncertainty, the dynamics of the economy, preferences and prices. This paper focuses on a derivatives-related quantity called the variance premium, which is measured as the difference between (the square of) the CBOE's VIX index and the conditional expectation of realized variance.

The variance premium provides a useful vehicle for understanding the connection between uncertainty in the economy and asset prices. An example of this connection, which we take as further motivation for focusing on the the variance premium, is the recent evidence presented in Bollerslev and Zhou (2007) highlighting the large, statistically significant predictive power of the variance premium for stock market returns. In this paper, we show theoretically that the variance premium is intimately linked to fundamental economic uncertainty and we derive conditions under which it predicts future stock returns. We further document several important new empirical features of the variance premium and its return predictability. We then proceed to analyze whether a generalized Long Run Risks (LRR) model with a rich set of transient dynamics can quantitatively account for the time variation and return predictability of the variance premium in addition to matching 'standard' asset pricing moments such as the level and volatility of the equity premium and risk free rate.

The variance premium is the difference between the price and expected payoff of a trading strategy that pays off best when the realized variance of returns is high. That this difference is essentially always positive (the price is higher than the expected payoff) reflects its function as a hedging strategy. Agents are willing to pay an insurance premium for an asset whose payoff is high when shocks to the economic state are large (i.e. when realized return variation is large). The conditional variance, a member of the state vector, is particulary volatile and prone to sharp increases. Moreover, volatility increases are associated with large negative returns, a well-documented phenomenon known as the 'leverage effect.' Hence, hedging shocks to volatility is an important function of the aforementioned variance asset. The high price of this asset (i.e. its role as a hedge) highlights the idea that agents must be averse to volatility/uncertainty in the state of the economy. In the Long Run Risks model of Bansal and Yaron (2004), agents prefer early resolution of uncertainty and consequently dislike economic uncertainty. Hence, variation in economic uncertainty, as captured by time variation of cashflow volatility (e.g., consumption and dividend growth), is, in that model, an important priced risk-source that leads to time varying risk premia.[1] This feature of the model also plays a prominent role in accounting for time variation in the variance premium. Specifically, we show that time variation in economic uncertainty and a preference for early resolution of uncertainty are required to generate a positive variance premium that is time varying and that predicts future excess returns.[2]

Our analysis shows that the compensation for economic uncertainty within the LRR model captures many qualitative features of the variance premium. However, we show that the baseline model requires several important modifications in order to capture quantitative features of the data such as the standard deviation and skewness of the variance premium, and its ability to predict stock returns. An extended specification of Bansal and Yaron (2004) which includes a square root volatility process and jumps in volatility goes a long way in capturing these quantitative aspects of the data. The inclusion of some non-gaussian transient dynamics is important for capturing key features of the variance premium; and it is important to note that the model is able to do so while being consistent with consumption and dividend dynamics, as well as standard asset prices such as the equity premium and risk free rate.

The paper continues as follows: Section 2 presents the data, defines the variance premium, discusses its statistical properties, and then proceeds to evaluate its role in predicting future returns. In section 3 we first present an illustration model which serves to demonstrate the link between the variance premium and return predictability. In Section 4 we present a more elaborate model that incorporates jumps in cashflows and uncertainty. Section 5 provides results from calibrating several specifications of these models. Section 6 provides concluding remarks.

[1]Bansal, Khatchatrian, and Yaron (2005) provide empirical evidence supporting the presence of conditional volatility in cashflows across several countries, Lettau, Ludvigson, and Wachter (2007) analyze whether the great moderation, the decline in aggregate volatility of macro aggregates can reconcile the run-up in valuation ratios during the late 90s. Bloom (2007) provides direct evidence linking spikes in market return uncertainty and subsequent decline in economic activity.

[2]Tauchen (2005) generalizes the volatility uncertainty in Bansal and Yaron (2004) to one with time varying economic uncertainty. Eraker (2007) adds jumps to the volatility specification. The focus on the variance premium is different from these papers.

2 Definitions and Data

Our definitions of key terms are similar to those in Bollerslev and Zhou (2007) and closely follow the related literature. We formally define the variance premium as the difference between the risk neutral and physical expectations of the market's total return variation. We will 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[Total Return Variation$(t, t+1)$] $-E_t$[Total Return Variation$(t, t+1)$]. 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. Thus, we obtain closing values of the VIX from the CBOE and use it as our measure of risk-neutral expected variance. Since the VIX index is reported in annualized "vol" terms, we square it to put it in "variance" space and divide by 12 to get a monthly quantity. Below we refer to the resulting series as squared VIX.

As the definition of $vp_{t,t+1}$ indicates, we also need conditional forecasts of total return variation under the true data generating process or physical measure. To obtain these forecasts we create measures of the total realized variation of the market, or realized variance, for the months in our sample. Our measure is created by summing the squares five-minute log returns over a whole month. For comparison, we do this for both the S&P 500 futures and S&P 500 cash index. We obtain the high frequency data used in the construction of our realized variance measures from TICKDATA. As discussed below, we project the realized variance measures on a set of predictor variables and construct forecasted series for realized variance. These forecast series are our proxy for the conditional expectation of total return variance under the physical measure. The difference between the risk neutral expectation, measured using the VIX, and the conditional forecasts from our projections, gives the series of one-month variance premium estimates.

Our data series covers the period January 1990 to March 2007. The main limitation on the length of our sample comes from the VIX, which is only published by the CBOE beginning in January of 1990. We use CRSP for daily and monthly returns on the value weighted return on the NYSE, AMEX and NASDAQ and the S&P 500. Finally, we obtain the monthly P/E ratio on the S&P 500 from Global Financial Data.

Table I provides summary statistics for the monthly log excess returns on both the S&P 500 and the value weighted return. The excess returns are constructed by subtracting the log 30-day T-Bill return, available from CRSP. The two series display very similar statistics. Both series have an approximately 0.53% mean monthly excess return with a volatility of about 4%. The other statistics are also quite close. Thus, although the availability of high-frequency data for the S&P 500 leads us to focus on it in our empirical analysis, our empirical inferences and theoretical model apply to the broader market.

The last four columns in Table I provide statistics for several measures of realized variance — potential inputs for our forecasts of realized variance: the squared VIX, the futures realized variance, cash index realized variance, and also the sum of squared daily returns over the month. The squared VIX value for a particular month is simply the value for the last day of the month. The futures, cash, and daily realized variances are sums over the whole month. We will ultimately use the futures realized variance and we display the other two for comparison. Several issues are worth noting. First, all volatility measures 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. Bollerslev and Zhou (2007) use the sum based on the cash index returns as their realized variance measure. This realized variance has a smaller mean than the futures and daily measures. This smaller mean is a result of a non-trivial autocorrelation in the five-minute returns on the cash index and is not present in the returns on the futures. We suspect that this autocorrelation is the effect of 'stale' prices at the five-minute intervals, since computation of the S&P 500 cash index involves 500 separate prices. As the S&P 500 futures involves only one price, and has long been one of the most liquid financial instruments available, we choose to use its realized variance measure to proxy for the total return variation of the market.

Table II provides a comparison of conditional variance projections. Our approach is to find a parsimonious representation, yet one that delivers significant predictability. The last two regressions show our choice of projection for the S&P index and futures variance measures. For these dependent variables we find that a parsimonious projection on the lagged VIX and index realized variance achieve adjusted R^2s of close to 60%. The addition of further lags or predictor variables adds very little predictive power. The first regression in the table provides the conditional volatility based on the daily squared returns. We fit a GARCH(1,1) to provide a comparison with the approaches used in early studies of variation, which used daily data. This regression also achieves an R^2 of around 40%. It is the use of high-frequency returns and the VIX as predictor that accomplishes the increased predictive power of the first two regressions.

Table III provides summary statistics for various measures of the variance premium, constructed as differences of the squared VIX and various variance forecasts. For comparison, the first column also reports the measure used by Bollerslev and Zhou (2007). They calculate the variance premium by subtracting from the squared VIX the previous month's realized variance. It is apparent from the table that the mean of the variance premium is somewhat smaller when based on the cash index measures as opposed to the futures or daily variance measures. Furthermore, the variance premium based on the futures measure is significantly less volatile than the other measures. Neither effects are surprising given the results in Table II and the discussion above regarding the cash index realized variance. The remaining statistics, in particular the skewness and kurtosis, seem to be quite similar across the variance premium proxies. In what follows, we use the variance premium based on the futures realized variance. As discussed above, the liquidity of the futures contract makes it an appropriate instrument for measuring realized variance. It is also the defacto instrument used by traders involved in related options trading. It is important to note however that our subsequent results are not materially effected by the use of this particular measure.

Table IV provides the one-month ahead return predictability regressions. In the first two regressions we use the variance premium as a univariate regressor. The latter specifications use multiple regressors where, in addition to the variance premium, we add common predictive variables, such as the price-dividend and price-earnings ratio. As a univariate regressor, the variance premium can account for about 3.5% of the return variation. The multivariate regressions lead to a substantial further increase in the adjusted R^2 – a feature highlighted in Bollerslev and Zhou (2007). For example, in conjunction with the price-earnings ratio, the in-sample adjusted R^2 increases to as much as 12.5%[3] It is worth noting that the lagged

[3] The in-sample adjusted R^2 of the price-earnings ratio alone is about 3.0%. The bivariate R^2s are significantly higher than the sum of R^2s from the univariate regressions. This is because of a positive correlation between the two regressors.

variance premium seems to perform better than the immediate variance premium. Note that in both cases, as well as the multivariate specification, the variance premium enters with a significant positive coefficient. Below, we show that this sign and magnitude are consistent with theory. A natural question that arises is whether such R^2s are economically significant. Cochrane (1999) uses a theorem of Hansen and Jagannathan (1991) to derive a relationship between the maximum unconditional Sharpe ratio attainable using a predictive regression and the regression R^2. It says that $(s^*)^2 - s_0^2 = \frac{1+s_0^2}{1-R^2}R^2$, where s_0 is the unconditional buy-and-hold Sharpe Ratio and s^* is the maximum unconditional Sharpe ratio[4]. In our sample, s_0 is approximately 0.157 at a monthly frequency, or 0.543 annualized. Using the univariate regression with an R^2 of 3.44%, the maximal Sharpe ratio would rise to 0.857 annualized. With the bivariate R^2 of 7.44%, the maximal Sharpe Ratio would further increase to 1.13, more than double the unconditional ratio. In other words, the potential increases are quite large. It is important to keep in mind that these R^2s are for a *monthly* horizon, and that Sharpe ratios increase roughly with the square root of the horizon. Hence an R^2 of 3% at the monthly horizon is potentially very useful. A comparison with "traditional" predictive variables found in the literature also shows this predictability is large. For example, Campbell, Lo, and MacKinlay (1997) examine the standard price-dividend ratio and stochastically detrended short-term interest rate, two of the more successful predictive variables, and show that in the more predictable second subsample, the predictive R^2s are 1.5% and 1.9% respectively at the monthly horizon. Campbell and Thompson (2007) examine a large collection of predictive variables whose in-sample (monthly) R^2s are much smaller than those reported in Table IV, but still conclude that these variables can be useful to investors. Finally, note that the variance related variables, i.e. the VIX^2, realized variance measures, and variance premium, all have AR(1) coefficients of 0.79 or less, unlike the price-dividend ratio or short term interest rate, which have AR(1) coefficients much closer to 1. This means the variance related quantities will not suffer from the large predictive regression biases associated with extremely persistent predictive variables, such as the price-dividend ratio (e.g. Stambaugh (1999)), and will have much better finite sample properties.

[4]This formula corresponds to the case when the predictive regression's residual is homoskedastic. If the predictive regressor also forecasts increased residual variance, the improvement in unconditional sharpe ratio will be less. This is clearly the case here since the predictors are closely related to volatility forecasts. Hence, we are *not* using the formula to draw any conclusions about attainable R^2s, but only to show that the R^2 sizes are economically meaningful

3 Model I

In this section we specify a model based on Bansal and Yaron (2004). The underlying environment is one with complete markets and the representative agent has Epstein and Zin (1989) type preferences which allow for the separation of risk aversion and the elasticity of intertemporal substitution. Specifically, the agent maximizes her life-time utility, which is defined recursively as,

$$V_t = \left[(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\Big(E_t\big[V_{t+1}^{1-\gamma}\big]\Big)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}} , \tag{1}$$

where C_t is consumption at time t, $0 < \delta < 1$ reflects the agent's time preferences, γ is the coefficient of risk aversion, $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, and ψ is the elasticity of intertemporal substitution (IES). Utility maximization is subject to the budget constraint,

$$W_{t+1} = (W_t - C_t)R_{c,t+1} , \tag{2}$$

where W_t is the wealth of the agent, and $R_{c,t}$ is the return on all invested wealth.

We provide several specifications for the joint dynamics of consumption and dividends. The specifications differ only along the the dynamics for the conditional volatility of consumption and dividends. We start with the following specification:

$$\begin{aligned}
\Delta c_{t+1} &= \mu_c + x_t + \varphi_c \sigma_t z_{c,t+1} \\
x_{t+1} &= \rho_x x_t + \varphi_x \sigma_t z_{x,t+1} \\
\Delta d_{t+1} &= \mu_d + \phi x_t + \varphi_d \sigma_t z_{d,t+1} \\
\sigma_{t+1}^2 &= \bar{\sigma}^2 + \rho_\sigma(\sigma_t^2 - \bar{\sigma}^2) + q_t^{1/2} z_{\sigma,t+1} && (3) \\
q_{t+1} &= \bar{q} + \rho_q(q_t - \bar{q}) + \varphi_q z_{q,t+1} && (4)
\end{aligned}$$

$$(z_{c,t+1}, z_{x,t+1}, z_{d,t+1}, z_{\sigma,t+1}, z_{q,t+1})' \sim \mathcal{N}(0, \mathcal{I})$$

where Δc_{t+1} is the growth rate of log consumption. As in the long run risks model of Bansal and Yaron (2004) (BY), $\mu_c + x_t$ is the conditional expectation of consumption growth, and x_t is a small but persistent component that captures long run risks in consumption growth. The parameter ρ_x determines the persistence in the conditional mean of consumption growth, $\mu_c + x_t$.

Our first specification for the conditional volatility has, in the interest of parsimony, a common time-varying volatility in consumption and dividends. Bansal and Yaron (2004) show that this leads to a time-varying risk premia. However, as we show below, that is not enough to get a time varying *variance premia*. To facilitate time variation of the variance premia, we specify a process for q_t, inducing time variation in the volatility process itself.

As shown in Epstein and Zin (1989), for any asset j, the first order condition yields the following asset pricing Euler condition,

$$E_t \left[\exp\left(m_{t+1} + r_{j,t+1}\right)\right] = 1 \tag{5}$$

where m_{t+1} is the log of the intertemporal marginal rate of substitution and $r_{j,t+1}$ is the log of the gross return on asset j. The solution for the log price-consumption assets is going to be driven by the three state variables x_t, σ_t^2, and q_t. As shown in Bansal and Yaron (2004), using this Euler equation, and as shown in Bansal and Yaron (2004), the approximate analytical no-bubble solutions for the log price-consumption asset follows:

$$v_t = A_0 + A_x x_t + A_\sigma \sigma_t^2 + A_q q_t$$

The solutions for As are given below and a detailed derivation is given in the Appendix:

$$A_x = \frac{1 - \frac{1}{\psi}}{1 - \kappa_1 \rho_x} \tag{6}$$

$$A_\sigma = \frac{1}{2}\theta \frac{\left[(1 - \frac{1}{\psi})^2 + (\kappa_1 A_x \varphi_x)^2\right]}{(1 - \kappa_1 \rho_\sigma)} \tag{7}$$

$$A_q = \frac{\theta(\kappa_1 A_\sigma)^2}{2(1 - \kappa_1 \rho_q)} \tag{8}$$

It is easily shown (see the appendix) that the (log of the) IMRS can be written in terms of the three state variables, the shocks, and the solutions to above as follows,

$$m_{t+1} - E_t(m_{t+1}) = -\lambda_c \varphi_c \sigma_t z_{c,t+1} - \lambda_x \varphi_x \sigma_t z_{x,t+1} - \lambda_\sigma q_t^{1/2} z_{\sigma,t+1} - \lambda_q \varphi_q z_{q,t+1} \tag{9}$$

where the market price of risks are $\lambda_c = \gamma$ for transient shocks, $\lambda_x = (1 - \theta)\kappa_1 A_x$ for the long run shocks, $\lambda_\sigma = (1-\theta)\kappa_1 A_\sigma$ for the economic uncertainty (volatility) shocks, and $\lambda_q = (1-\theta)\kappa_1 A_q$ for the shocks to volatility of volatility. Note that the dividend-specific shock, $z_{d,t+1}$, is not priced, i.e. $\lambda_d = 0$. The return on the consumption asset will be a function of these same state variables (see equation **??** in the appendix). It is important to recognize that when the preferences are specialized to time separable CRRA, namely $\theta = 1$, then only the transient risk is priced while all others are zero, i.e. $\lambda_x = \lambda_\sigma = \lambda_q = 0$. Moreover, if $\psi > 1$, the market price of volatility is *negative*, a fact that will have important ramifications for the variance premium.

The solution for the log price dividend ratio for the market will follow $v_{m,t} = A_{0,m} + A_{x,m}x_t + A_{\sigma,m}\sigma_t^2 + A_{q,m}q_t$. The solutions for the A_m coefficients can be derived by specializing the Euler equation (4.1) for the market return and these are given in the appendix —- equation (**??**).

From the expression for $v_{m,t}$ it follows that the innovation to the (linearized) log return on the market can be written as

$$r_{m,t+1} - E_t(r_{m,t+1}) = \underbrace{\kappa_{1,m}A_{x,m}\varphi_x}_{\beta_{r,x}}\sigma_t z_{x,t+1} + \underbrace{\kappa_{1,m}A_{\sigma,m}}_{\beta_{r,\sigma}}q_t^{1/2}z_{q,t+1} + \underbrace{\kappa_{1,m}A_{q,m}\varphi_q}_{\beta_{r,q}}z_{q,t+1} + \varphi_d\sigma_t z_{d,t+1}$$

It immediately follows that the conditional variance of the market return is

$$\text{var}_t(r_{m,t+1}) = \sigma_{r,t}^2 = (\beta_{r,x}^2 + \varphi_d^2)\sigma_t^2 + \beta_{r,\sigma}^2 q_t + \beta_{r,q}^2 \tag{10}$$

Much of our focus will be on a concept defined as the *variance premium*, which is measured using realized volatility. The empirical literature on realized volatility usually computes realized variance using the (log) ex-dividend return, i.e. the capital gain or (log) price change. We follow the same convention within the model and compute the volatility of the (log) price change, Δp_{t+1}, which is equal to $v_{m,t+1} - v_{m,t} + \Delta d_{t+1}$. Since this is expression is very similar to the one for the total return, its conditional variance has the same structure as (10) and can be conveniently written

$$\text{var}_t(\Delta p_{t+1}) = \sigma_{p,t}^2 = (\beta_{p,x}^2 + \varphi_d^2)\sigma_t^2 + \beta_{p,\sigma}^2 q_t + \beta_{p,q}^2 \tag{11}$$

where now the β's (described in the appendix) have a p subscript (instead of an r as before) to reflect their relation to the *p*rice gain component of the total return[5].

3.1 The Variance Premium

The variance premium, denoted by vp, is the difference between the rep agent's expectation of the integral of variance under the risk neutral measure (Q) and the physical measure (P). In a discrete-time model, with n sub-periods between t and $t+1$, the integral in the definition of the variance premium is replaced with a sum and the variance premium is defined as follows:

$$vp_{t,t+1} \equiv E_t^Q[\sum_{i=1}^{n-1} \sigma^2_{p,t+\frac{i-1}{n},t+\frac{i}{n}}] - E^P[\sum_{i=1}^{n-1} \sigma^2_{p,t+\frac{i-1}{n},t+\frac{i}{n}}] \tag{12}$$

where $E_t^Q(\cdot)$ denotes the expectation under the risk-neutral measure and where $\sigma^2_{p,t+\frac{i-1}{n},t+\frac{i}{n}}$ denotes the quantity $\mathrm{var}_{t+\frac{i}{n}}(\Delta p_{t+\frac{i-1}{n},t+\frac{i}{n}})$ when it appears in $E^P(\cdot)$ and the quantity $\mathrm{var}^Q_{t+\frac{i}{n}}(\Delta p_{t+\frac{i-1}{n},t+\frac{i}{n}})$ when it appears in $E^Q(\cdot)$. Since the variance premium involves expectations under Q of functions of the state variables, we need the dynamics of the state variables under Q.

3.1.1 Dynamics under Q

Under the physical measure, the dynamics of the model can be concisely expressed as a VAR,

$$y_{t+1} = c + Fy_t + G_t z_{t+1}$$

where $y_{t+1} = (\Delta c_{t+1}, x_{t+1}, \Delta d_{t+1}, \sigma^2_{t+1}, q_{t+1})'$, $c = (\mu_c, 0, \mu_d, \bar{\sigma}^2(1-\rho_\sigma), \bar{q})'$, G_t is the diagonal matrix with diagonal equal to $(\varphi_c\sigma_t, \varphi_x\sigma_t, \varphi_d\sigma_t, q_t^{1/2}, \varphi_q)'$ and z_{t+1} is the vector of independent normal innovations. Let $f_t(z_{t+1})$ denote the joint density of z_{t+1} at time t under the physical measure. Then $f_t(z_{t+1}) \propto \exp(-\frac{1}{2}z'_{t+1}z_{t+1})$. To obtain $f_t^Q(z_{t+1})$, the joint density under Q, we weight the density by the value of the pricing kernel, $M_{t+1}(z_{t+1}) = \exp(m_{t+1}(z_{t+1}))$. From (9) we see that $m_{t+1}(z_{t+1}) = E_t(m_{t+1}) - \lambda' G_t z_{t+1}$ where λ is the vector of risk prices

[5]For all practical purposes, both empirically and in any calibrations, $\sigma^2_{r,t}$ and $\sigma^2_{p,t}$ are virtually identical as are all the pairs corresponding β_r's, and β_p's

$(\lambda_c, \lambda_x, \lambda_d, \lambda_\sigma, \lambda_q)'$. Therefore,

$$\begin{aligned} f_t^Q(z_{t+1}) &\propto f_t(z_{t+1}) \exp(m(z_{t+1})) \\ &\propto \exp(-\frac{1}{2} z'_{t+1} z_{t+1}) \exp(-\lambda' G_t z_{t+1}) \\ &\propto \exp(-\frac{1}{2}(z_{t+1} + G'_t \lambda)'(z_{t+1} + G'_t \lambda)) \end{aligned}$$

where the last line follows from the familiar "complete-the-square" argument. This shows that $z_{t+1} \overset{Q}{\sim} \mathcal{N}(-G'_t\lambda, I)$, i.e. under Q, z_{t+1} are still independent normals but with a shifted mean. Thus, under Q the VAR's dynamics are $y_{t+1} = c + Fy_t - G_t G'_t \lambda + G_t \tilde{z}_{t+1}$ where now $\tilde{z}_{t+1} \overset{Q}{\sim} \mathcal{N}(0, I)$. Therefore, under Q the "drift" of y is shifted by $-G_t G'_t \lambda$. More intuitively, $-G_t G'_t \lambda = \text{cov}_t(y_{t+1}, m_{t+1})$, so we see that

$$E_t^Q[y_{t+1}] - E_t[y_{t+1}] = \text{cov}_t(y_{t+1}, m_{t+1}) \tag{13}$$

i.e. the shift in the drift of y_t equals the conditional covariance of y_{t+1} with the pricing kernel.

Although the drift of y is altered under the risk-neutral measure, the innovation to y remains $\mathcal{N}(0, G_t G'_t)$. This means that any quantity that is a function of only the innovation to y_{t+1} will retain the same functional form under Q. In particular, $\text{var}_t^Q(r_{m,t+1})$ is still given by (10) and $\text{var}_t^Q(\Delta p_{t+1})$ by (11). Thus, the term $\sigma^2_{p,t}$ represents an unambiguous function of the state y_t that does *not* depend on which measure is under consideration. This invariance of the expression for the conditional variance results from the gaussian nature of the innovations in this model. Later, we discuss a model with some non-gaussian innovations where the conditional variance will differ between the P and Q measures.

Returning to the expression for $vp_{t,t+1}$ in (12), we now have that the terms $E[\sigma^2_{p,t,t+\frac{1}{n}}]$ and $E^Q[\sigma^2_{p,t,t+\frac{1}{n}}]$ (the first summands) are equivalent, since $\sigma^2_{p,t,t+\frac{1}{n}}$ is in the time-t information set and is an identical function of y_t under both measures. However, for $i \geq 1$, $E[\sigma^2_{p,t+\frac{i-1}{n},t+\frac{i}{n}}] \neq E^Q[\sigma^2_{p,t+\frac{i-1}{n},t+\frac{i}{n}}]$ since the drift of y is different under Q, and so $\sigma^2_{p,t+\frac{i-1}{n},t+\frac{i}{n}}$ drifts differently under the two measures. Thus, under this model, the variance premium is non-zero only because the variances drift differently. [6] When we look at a non-gaussian model later on,

[6] This result is the discrete-time analog to what is typically the case in continuous-time diffusion models of option pricing, though it is perhaps less obvious under the continuous-time formulations. For example, in the well-known Heston (1993) model, the variance premium for the "dt" interval [t,t+dt) is actually 0. It is non-zero for any finite interval $[t, t + \delta t)$ precisely because of the differential drift in σ^2_t under Q and P.

the variance premium will also get a significant contribution from the difference between $\text{var}_t^Q(\Delta p_{t+1})$ and $\text{var}_t(\Delta p_{t+1})$.

Therefore, under the current model, the variance premium can be understood by comparing the drift of $\sigma_{p,t}^2$ under Q and P, i.e. by deriving

$$E^Q[\sigma_{p,t+1}^2 - \sigma_{p,t}^2] - E[\sigma_{p,t+1}^2 - \sigma_{p,t}^2] = E^Q[\sigma_{p,t+1}^2] - E[\sigma_{p,t+1}^2]$$

The expression for $vp_{t,t+1}$ in (12) is effectively a sum of n of these drift differences over the sub-periods from $t + \frac{i-1}{n}$ to $t + \frac{i}{n}$. For simplicity, and since the same principle is at work, in what follows we approximate $vp_{t,t+1}$ with just $E^Q[\sigma_{p,t+1}^2] - E[\sigma_{p,t+1}^2]$, the drift difference over the whole interval t to $t+1$, and use this drift difference as the new definition of $vp_{t,t+1}$.

3.1.2 The Drift Difference

Specializing (13) to just the state variables σ_t^2 and q_t, and using (3), (4), and (9), an easy calculation gives

$$E_t^Q[\sigma_{t+1}^2] - E_t[\sigma_{t+1}^2] = -\lambda_\sigma q_t \tag{14}$$

$$E_t^Q[q_{t+1}] - E_t[q_{t+1}] = -\lambda_q \varphi_q^2 \tag{15}$$

Recall that $\gamma, \psi > 1$ implies that $\lambda_\sigma, \lambda_q < 0$ so that the drift in σ_t^2 and q_t is *greater* under the risk-neutral measure than the physical measure. The intuition is straightforward. Under the risk-neutral measure, the agent weights the true probabilities of states of the world by his IMRS (the pricing kernel). Under Epstein-Zin preferences, innovations to the level of uncertainty, σ_t^2, impact the IMRS. The degree to which they impact the IMRS is governed by λ_σ, the price of volatility risk. When $\gamma, \psi > 1$, then $\lambda_\sigma < 0$ and a positive innovation to uncertainty increases m_{t+1} (see (9)), i.e. the agent over-weights states with high volatility since he is averse them. This implies that the expectation for σ_{t+1}^2 under Q should exceed the expectation under P. The amount of this excess (at time t) depends on the possible variability in σ_{t+1}^2, that is, on q_t. For the drift in q_t the same reasoning applies since the agent's aversion to high σ_{t+1}^2 leads to aversion of high q_{t+1}. As previously noted, under CRRA preferences, $\lambda_\sigma = \lambda_q = 0$, so the drift differences (14), (15) would be 0 as well.

From (11), (14), and (15), it easily follows that, $E^Q[\sigma_{p,t+1}^2] - E[\sigma_{p,t+1}^2] = -(\beta_{p,x}^2 + \varphi_d^2)\lambda_\sigma q_t -$

$\beta_{p,\sigma}^2\lambda_q\varphi_q^2$, so finally we obtain

$$vp_{t,t+1} = -(\beta_{p,x}^2 + \varphi_d^2)\lambda_\sigma q_t - \beta_{p,\sigma}^2\lambda_q\varphi_q^2 \tag{16}$$

Some observations regarding (16): (1) If $\gamma, \psi > 1$ then $\lambda_\sigma < 0$ as noted above, and the variance premium is always *positive*, consistent with the data. (2) In the CRRA case ($\psi = \frac{1}{\gamma}$), the variance premium is identically 0. (3) If $\gamma > 1$ but $\psi < \frac{1}{\gamma}$, then the variance premium is always *negative*. (4) Time variation in $vp_{t,t+1}$ is due to variation in q_t so if q_t is constant, as is the case in the specification of Bansal and Yaron (2004), then the variance premium is constant as well.

3.2 Predictability

Motivated by the predictability results in section 2 and Bollerslev and Zhou (2007) we now present analytical projection coefficients and R^2s for predictive regressions of future returns on the variance premium. Let β_{vp} be defined by the following projection:

$$r_{m,t+1} - r_{f,t} + 0.5\sigma_{r,t}^2 = \alpha + \beta_{vp} vp_{t,t+1} + \epsilon_{t+1}$$

It follows that,

$$\begin{aligned} \beta_{vp} &= \frac{cov(E_t(r_{m+1} - r_{f,t}) + 0.5\sigma_{m,t}^2 + \tilde{\epsilon}_{t+1}, vp_{t,t+1})}{\text{var}(vp_{t,t+1})} \\ &= \frac{cov(\beta_{r,x}\lambda_x\varphi_x\sigma_t^2 + \beta_{r,\sigma}\lambda_\sigma q_t + \beta_{r,q}\lambda_q\varphi_q, -(\beta_{p,x}^2 + \varphi_d^2)\lambda_\sigma q_t - \beta_{p,\sigma}^2\lambda_q\varphi_q^2)}{\text{var}(vp_{t,t+1})} \\ &= \frac{-\beta_{r,\sigma}}{\beta_{p,x}^2 + \varphi_d^2} \end{aligned} \tag{17}$$

where the second line makes use of the orthogonality of $\tilde{\epsilon}_{t+1}$ to $vp_{t,t+1}$ and substitutes in the expression for the market risk premium [Note: we need to derive/introduce the expression for the risk premia], while the third line uses the fact that $\text{cov}(q_t, \sigma_t) = 0$[7]. We note again when $\gamma > 1, \psi > 1$ the sign of β_{vp} is positive, as in the data, since in that case $\beta_{r,\sigma}$ is negative

[7] Note that in deriving β_{vp} above, we added the Jensen inequality term ($\sigma_{r,t}^2$) to the risk premia. Including the term makes the above projection similar to one that predicts simple rather than log returns. Empirically, the two yields very similar results. We include this term above to allow for simple analytical expressions that provide useful intuition for the underlying sources of covariation between $vp_{t,t+1}$ and risk premia.

[8].

Finally, we derive the R^2 of the above projection. It is $\frac{\beta_{vp}^2 \text{var}(vp_{t,t+1})}{\text{var}(r_{m,t+1} - r_{f,t} + 0.5\sigma_{r,t}^2)}$. Substituting into the numerator and simplifying gives $(\beta_{r,\sigma}\lambda_\sigma)^2\text{var}(q_t)$. The denominator can be derived using the identity $\text{var}(X) = E[\text{var}_t(X)] + \text{var}[E_t(X)]$. Since $\text{var}_t(r_{m,t+1} - r_{f,t} + 0.5\sigma_{r,t}^2) = \text{var}_t(r_{m,t+1})$ we can use (10). After some substitutions one obtains

$$R^2 = \frac{(\beta_{r,\sigma}\lambda_\sigma)^2\text{var}(q_t)}{(\beta_{r,x}^2 + \varphi_d^2)\bar{\sigma}^2 + \beta_{r,\sigma}^2\bar{q} + \beta_{r,q}^2 + (\beta_{r,x}\lambda_x\varphi_x)^2\text{var}(\sigma_t^2) + (\beta_{r,\sigma}\lambda_\sigma)^2\text{var}(q_t)} \tag{18}$$

(18) highlights that, under this model, an increase in $\text{var}(q_t)$ (equivalently $vp_{t,t+1}$) with all else held constant, raises the R^2 of the predictive regression. This variance is controlled by the parameter φ_q. Finally, note that the larger the magnitude of λ_σ, the greater is the effect of q_t on the conditional risk premia. This is because λ_σ measures the importance of σ_{t+1}^2 innovations for the agent's IMRS, and q_t drives the magnitude of those innovations.

4 Model with Jumps

The state variables follow the discrete-time affine jump-diffusion

$$\Delta c_{t+1} = \mu_c + x_t + \varphi_c\sigma_t^c z_{c,t+1} \tag{19}$$

$$x_{t+1} = \rho_x x_t + \varphi_x\sigma_t z_{x,t+1} + J_{x,t+1} \tag{20}$$

$$\sigma_{t+1}^2 = \bar{\sigma}^2 + \rho_\sigma(\sigma_t^2 - \bar{\sigma}^2) + \varphi_\sigma\sigma_t z_{\sigma,t+1} + J_{\sigma,t+1} \tag{21}$$

$$\Delta d_{t+1} = \mu_d + \phi x_t + \varphi_d\sigma_t^d z_{d,t+1} \tag{22}$$

$$J_{x,t+1} = \sum_{i=1}^{N_{t+1}^x} \xi_{x,i} \qquad \xi_{x,i} \overset{iid}{\sim} \mathcal{N}(0, \sigma_x^2)$$

$$J_{\sigma,t+1} = \sum_{i=1}^{N_{t+1}^\sigma} \xi_{\sigma,i} \qquad \xi_{\sigma,i} \overset{iid}{\sim} \Gamma(\nu, \frac{\mu_\sigma}{\nu})$$

and $(z_{c,t+1}, z_{x,t+1}, z_{\sigma,t+1}, z_{d,t+1}) \sim \mathcal{N}(0, \mathcal{I})$. N_{t+1}^x and N_{t+1}^σ are Poisson counting processes that are independent conditional on time-t information. $\xi_{x,i}$ and $\xi_{\sigma,i}$ are, respectively, the

[8] The sign and magnitude of $\beta_{r,\sigma}$ is mostly determined by the term $(1-\theta)A_\sigma$, which is itself almost the definition of λ_σ. As discussed earlier, λ_σ is negative for $\gamma, \psi > 1$, 0 for CRRA, and positive when $\gamma > 1$, $\psi < \frac{1}{\gamma}$.

jumps in x and σ that occur when the respective Poisson process increments. Hence, $J_{x,t+1}$ and $J_{\sigma,t+1}$ are compound Poisson processes representing the jumps in x and σ between time t and $t+1$. To limit the number of state variables, we let N^x_{t+1} and N^σ_{t+1} share a single jump intensity process, λ_t, and make this intensity depend only on σ^2_t:

$$\lambda_t = \lambda_0 + \lambda_1 \sigma^2_t \qquad \lambda_0, \lambda_1 > 0$$

Note that ξ_σ has a gamma distribution with $E[\xi_\sigma] = \mu_\sigma$, $\text{var}(\xi_\sigma) = \frac{\mu^2_\sigma}{\nu}$, skewness $= \frac{2}{\sqrt{\nu}}$ and excess kurtosis $= \frac{6}{\nu}$. This type of jump is used in Eraker and Shaliastovich (2007) and it subsumes the exponential jumps used in Eraker (2007) and Duffie, Pan, and Singleton (2000). Since N^σ_{t+1} is Poisson with intensity λ_t, $E_t[J_{\sigma,t+1}] = \lambda_t \mu_\sigma$. Therefore, changes in the the jump intensity *do* change the expected dynamics of σ^2_{t+1}. As x jumps have 0 mean, $E_t[J_{x,t+1}] = 0$, they do *not* change the expected dyanmics of x_{t+1}. It is helpful to rewrite the dynamics of σ^2_t with a conditionally de-meaned jump (i.e. a compensated Poisson process). The de-meaned jump is then a true 'innovation'. Let $\tilde{J}_{\sigma,t+1} = J_{\sigma,t+1} - \lambda_t \mu_\sigma$. We break up $\lambda_t \mu_\sigma$ into its two parts, $\lambda_1 \mu_\sigma \sigma^2_t$ and $\lambda_0 \mu_\sigma$ and define $\tilde{\rho}_\sigma = \rho_\sigma + \lambda_1 \mu_\sigma$. Then the dynamics of σ^2_{t+1} can be rewritten as

$$\sigma^2_{t+1} = \tilde{\bar{\sigma}}^2 + \tilde{\rho}_\sigma(\sigma^2_t - \tilde{\bar{\sigma}}^2) + \varphi_\sigma \sigma_t z_{\sigma,t+1} + \tilde{J}_{\sigma,t+1} \tag{23}$$

where $\tilde{\bar{\sigma}}^2 = \frac{\bar{\sigma}^2(1-\rho_\sigma) + \lambda_0 \mu_\sigma}{1 - \rho_\sigma - \lambda_1 \mu_\sigma} = E[\sigma^2_t]$. Note that $\tilde{\rho}$ is the first autocorrelation and must be less than 1 for σ^2_t to be stationary.

Finally, we define σ^c_t, the vol term multiplying consumption innovations, by $(\sigma^c_t)^2 = \tilde{\bar{\sigma}}^2(1-w_c) + \sigma^2_t$, and σ^d_t, for dividend innovations, by $(\sigma^d_t)^2 = \tilde{\bar{\sigma}}^2(1-w_d) + \sigma^2_t$. Thus, w_c and w_d control how sensitive the variance of consumption and dividend innovations is to σ^2_t. If $w_c = w_d = 1$, the innovation variances are fully sensitive, as in Bansal and Yaron (2004). On the other hand, if $w_c = w_d = 0$, then the innovation variances have no sensitivity to σ^2_t. Note that, changing the variance sensitivity has no effect on the unconditional mean of innovation variance, i.e. $E(\sigma^c_t)^2 = E(\sigma^d_t)^2 = \tilde{\bar{\sigma}}^2$.

For notational brevity in what follows, we also compactly write the dynamics of the model as a VAR,

$$y_{t+1} = \mu + F y_t + G_t z_{t+1} + J_{t+1} \tag{24}$$

where $y_{t+1} = (\Delta c_{t+1}, x_{t+1}, \sigma^2_{t+1}, \Delta d_{t+1})'$, $\mu = (\mu_c, 0, \bar{\sigma}^2(1-\rho_\sigma), \mu_d)$, G_t is a diagonal matrix with $(\varphi_c\sigma_t^c, \varphi_x\sigma_t, \varphi_\sigma\sigma_t, \varphi_d\sigma_t^d)$ on the diagonal, J_{t+1} is the vector of compound Poisson jumps, $(0, J_{x,t+1}, J_{\sigma,t+1}, 0)$, and z_{t+1} is the vector of normal innovations.

The specification of preferences is the same as above. In the following section we solve the model. The procedure is similar to the one above, albeit with some additional complexity.

4.1 Model Solution

The solution again proceeds via the rep agent's Euler condition (4.1). We conjecture and verify that the approximate, analytical no-bubbles solution for the log price-consumption ratio takes the form:

$$v_t = A_0 + A'y_t$$

where $A = (A_c, A_x, A_\sigma, A_d)'$ is a vector of pricing coefficients. The derivation of A is in the appendix. Here we make some remarks regarding key steps in the solution. It starts with a central step involved in the solution of affine-jump diffusions. This step is to show that for any (4x1) vector $\mathbf{u}$ meeting some regularity criteria, we have the following identity

$$E[\exp(\mathbf{u}'y_{t+1})|y_t] = \exp(\alpha(\mathbf{u}) + \beta(\mathbf{u})'y_t) \tag{25}$$

where $\alpha(\mathbf{u})$ is a scalar and $\beta(\mathbf{u})$ is a (4x1) vector. Once $\alpha(\mathbf{u})$ and $\beta(\mathbf{u})$ are known, (25) is used to evaluate the conditional expectation in the rep agent's Euler equation, . The vector $\mathbf{u}$ is directly a function of A, so it is actually a restriction on A. This restriction requires that the Euler equation hold for any value of the state vector y_t. By the method of undetermined coefficients, the restriction equates to a system of equations in A and A_0. Solving this system of equations completes the derivation of A and A_0.

There are similarities with the solution to model 1. As was the case for model 1, the solution shows that $A_c = A_d = 0$. This is not surprising. As in model 1, x_t subsumes all pricing information contained in Δc_{t+1} and Δd_{t+1}. Furthermore, $A_x = \dfrac{1-\frac{1}{\psi}}{1-\kappa_1\rho_x}$, also the same expression as under model 1. On the other hand, the case of A_σ is different, and no longer do we necessarily get a closed-form expression. Instead, A_σ must satisfy the following equation:

$$aA_\sigma^2 + bA_\sigma + c + \lambda_1\big(M_{\xi_\sigma}(A_\sigma\theta\kappa_1) - 1\big) = 0 \tag{26}$$

where $a = \frac{1}{2}(\varphi_\sigma\theta\kappa_1)^2$, $b = \theta(\rho_\sigma\kappa_1 - 1)$, and $c = \frac{1}{2}\big[\varphi_c^2 w_c\theta^2(1-\frac{1}{\psi})^2 + \varphi_x^2(A_x\theta\kappa_1)^2\big] +$

$\lambda_1\big(M_{\xi_x}(A_x\theta\kappa_1) - 1\big)$. The M_ξ's are the moment generating functions (mgfs) for ξ_σ and ξ_x. Hence, $M_{\xi_x}(t) = E[\exp(t\xi_x)] = \exp(\frac{1}{2}t^2\sigma_x^2)$ and $M_{\xi_\sigma}(t) = E[\exp(t\xi_\sigma)] = (1 - \frac{\mu_\sigma}{\nu}t)^{-\nu}$ [9]. Notice that if eliminate the jumps in volatility by setting $\mu_\sigma = 0$, then $M_{\xi_\sigma(t)} = 1$ and the last term in (26) drops out. In this case σ_t^2 reduces to a pure square-root process without jumps. A_σ then satisfies a quadratic equation and the nonexplosive root can be found analytically according to the procedure in Tauchen (2005). In the case where we include jumps, the form of the equation for A_σ depends on ν, the shape parameter of the gamma distribution. As ν decreases, the right tail becomes thicker and the distribution becomes more asymmetric. When $\nu = 1$, the gamma distribution reduces to an exponential distribution with mean μ_σ. In this special case, it is easy to show that (26) becomes a cubic in A_σ, and once more a closed-form expression can be found for the nonexplosive root. Moreover, this root converges to the nonexplosive quadratic root as $\mu_\sigma \to 0$ and jumps are eliminated. For $\nu < 1$, where the right tail is fatter than in the exponential case, there is no closed-form solution. The nonexplosive root must be found numerically. The solution procedure is discussed further in the appendix.

Having solved for the log price-consumption ratio, v_t, we can use it to write down the linearized (log of) the IMRS in terms of the state variables, their shocks and the A coefficients:

$$m_{t+1} - E_t(m_{t+1}) = -\lambda'\big(G_t z_{t+1} + J_{t+1} - E_t(J_{t+1})\big) = -\lambda'(G_t z_{t+1} + \tilde{J}_{t+1}) \tag{27}$$

where $\tilde{J}_{t+1} = (0, J_{x,t+1}, \tilde{J}_{\sigma,t+1}, 0)$ is the vector of compensated (i.e. conditionally demeaned) Poisson jumps and $\lambda = (\lambda_c, \lambda_x, \lambda_\sigma, \lambda_d)$ is the vector of prices of risk. The functional forms of these prices of risk are unchanged from model 1: $\lambda_c = \gamma$, $\lambda_x = (1-\theta)\kappa_1 2A_x$, $\lambda_\sigma = (1-\theta)\kappa_1 A_\sigma$, and $\lambda_d = 0$. The only change in the form of (27) from (9) is that the jump innovations are now added to the gaussian ones to get the total innovation. Moreover, the discussion above regarding the connection between preferences and the prices of risk carries over to this setting. It is again the case that when preferences are specialized to the CRRA case, only the transient risk, $z_{c,t+1}$, is priced and all other risk prices are 0. When $\gamma > 1, \psi > 1$, then again $\lambda_\sigma < 0$[10]. Finally, if $\gamma > 1$ but $\psi < \frac{1}{\gamma}$ then $\lambda > 0$ and volatility has a *positive* price of risk. As will be discussed below, this positive price of volatility risk would result in a counter-factually *negative* variance premium.

[9] We must have $t < \nu/\mu_\sigma$ for existence of the gamma mgf.

[10] Although A_σ does not necessarily have a closed-form expression it is still possible sign it in this case.

After solving for the consumption claim, one can proceed to solve for the market. The solution for the log price-dividend ratio of the market is again affine in the state variables: $v_{m,t} = A_{0,m} + A_{x,m}x_t + A_{\sigma,m}\sigma_t^2$. Solving for the A_m coefficients proceeds in much the same way as solving for the A coefficients of the log price-consumption claim. The Euler equation (4.1) is specialized to the market return $r_{m,t+1}$, leading to a system of equations that the A_m's must satisfy. The full details are in the appendix.

Using the expression for $v_{m,t}$, the (linearized) log return on the market can be written as

$$r_{m,t+1} - E_t(r_{m,t+1}) = \kappa_{1,m}A_{x,m}(\varphi_x\sigma_t z_{x,t+1} + J_{x,t+1}) + \kappa_{1,m}A_{\sigma,m}(\varphi_\sigma\sigma_t z_{\sigma,t+1} + \tilde{J}_{\sigma,t+1}) + \varphi_d\sigma_t^d z_{d,t+1}$$

where the parentheses group each state variable's total innovation. Since, conditional on time-t information, the jumps are independent from each other and also from the gaussian innovations, it follows that[11]

$$\mathrm{var}_t(r_{m,t+1}) = \sigma_{r,t}^2 = \beta_{\sigma_r}\sigma_t^2 + \beta_{\lambda_r}\lambda_t + \varphi_d^2\,\tilde{\bar{\sigma}}^2(1 - w_d) \tag{28}$$

where $\beta_{\sigma_r} = \kappa_{1,m}^2\left(A_{x,m}^2\varphi_x^2 + A_{\sigma,m}^2\varphi_\sigma^2\right) + \varphi_d^2 w_d$ and $\beta_{\lambda_r} = \kappa_{1,m}^2\left(A_{x,m}^2\sigma_x^2 + A_{\sigma,m}^2\left[\frac{\mu_\sigma^2}{\nu} + \mu_\sigma^2\right]\right)$. Note that although λ_t, the jump intensity, is not an independent state variable, we keep it separate to highlight its effects as distinct from the diffusion component. As discussed for model 1, we also have the corresponding expressions for the price change (capital gain), Δp_{t+1}, used to construct the empirical variance measures:

$$\mathrm{var}_t(\Delta p_{t+1}) = \sigma_{p,t}^2 = \beta_{\sigma_p}\sigma_t^2 + \beta_{\lambda_p}\lambda_t + \varphi_d^2\,\tilde{\bar{\sigma}}^2(1 - w_d) \tag{29}$$

where $\beta_{\sigma_p} = \left(A_{x,m}^2\varphi_x^2 + A_{\sigma,m}^2\varphi_\sigma^2\right) + \varphi_d^2 w_d$ and $\beta_{\lambda_p} = \left(A_{x,m}^2\sigma_x^2 + A_{\sigma,m}^2\left[\frac{\mu_\sigma^2}{\nu} + \mu_\sigma^2\right]\right)$.

We now obtain the dynamics of the model under Q, in order to derive the variance premium.

[11] Two identities useful in working with the jump components are: $\mathrm{var}_t(J_{x,t+1}) = \lambda_t\sigma_x^2$ and $\mathrm{var}_t(J_{\sigma,t+1}) = \lambda_t\left(\frac{\mu_\sigma^2}{\nu} + \mu_\sigma^2\right)$. These can be derived by using $\mathrm{var}_t(J_{t+1}) = E_t(\mathrm{var}_t(J_{t+1})|N_{t+1}) + \mathrm{var}_t(E_t(J_{t+1}|N_{t+1}))$ where the conditioning is on N_{t+1}, the realization of the Poisson counting process.

4.2 Variance Premium II

Since the variance premium involves an expectation under risk-neutrality, we need to derive the model dynamics under the risk neutral measure, Q.

Recall from (24) that under the physical measure the state dynamics can be written as the following VAR:

$$y_{t+1} = \mu + F y_t + G_t z_{t+1} + J_{t+1}$$

Moving from the physical to the risk-neutral measure transforms the state dynamics into:

$$y_{t+1} = \mu + F y_t - G_t G_t' \lambda + G_t \tilde{z}_{t+1} + J_{t+1}^Q \tag{30}$$

where $\tilde{z}_{t+1} \overset{Q}{\sim} \mathcal{N}(0, \mathcal{I})$ and J_{t+1}^Q is the vector of compound Poisson processes under Q. Under Q the compound Poisson processes remain conditionally independent, but *both* their jump intensities and jump distributions are different than under P. In fact, under Q, the two Poisson counting processes no longer share the same jump intensity. Let $\lambda_t^{Q,x}$ and $\lambda_t^{Q,\sigma}$ be the intensities of N_t^x and N_t^σ respectively, under Q. Then

$$\lambda_t^{Q,x} = \lambda_t M_{\xi_x}(-\lambda_x) = \lambda_t \exp\left(\frac{1}{2}\lambda_x^2 \sigma_x^2\right) \tag{31}$$

$$\lambda_t^{Q,\sigma} = \lambda_t M_{\xi_\sigma}(-\lambda_\sigma) = \lambda_t \left(1 + \frac{\mu_\sigma}{\nu}\lambda_\sigma\right)^{-\nu} \tag{32}$$

In addition, under Q the jump distributions of x and σ are transformed into the following:

$$\xi_{x,i} \overset{Q}{\sim} \mathcal{N}(-\lambda_x \sigma_x^2, \sigma_x^2) \tag{33}$$

$$\xi_{\sigma,i} \overset{Q}{\sim} \Gamma\left(\nu, \frac{\frac{\mu_\sigma}{\nu}}{1 + \frac{\mu_\sigma}{\nu}\lambda_\sigma}\right) \tag{34}$$

We make a number of observations regarding (31)–(34). Consider what happens to the jump intensities and distributions when $\lambda_x > 0$ and $\lambda_\sigma < 0$, which is the result when $\gamma > 1$, $\psi > 1$. Looking at the x quantities shows that $\lambda_t^{Q,x} > \lambda_t$ and $E^Q[\xi_{x,i}] < E[\xi_{x,i}] = 0$.[12] In other words, under the risk-neutral probabilities the agent expects more frequent jumps than under the physical measure. The agent also expects these jumps to be more negative. Both

[12]That $\lambda_t^{Q,x} > \lambda_t$ here does not depend on $\lambda_x > 0$. It is purely a Jensen's inequality effect that results because $\lambda_x \neq 0$. If $\xi_{x,i}$ had a mean different from 0, then the sign of λ_x would also be important in determining $\lambda_t^{Q,x}$.

effects imply that under Q the agent expects more variation in x_t than under P. Taking a look at the σ related quantities, we see that $\lambda_\sigma < 0$ implies $\lambda_t^{Q,\sigma} > \lambda_t$, so the agent expects more frequent jumps in σ_t^2 as well. Furthermore, $\lambda_\sigma < 0$ leads to an increase in the scale parameter of σ's jump distribution so that $E^Q[\xi_{\sigma,i}] = \dfrac{\mu_\sigma}{1+\frac{\mu_\sigma}{\nu}\lambda_\sigma} > \mu_\sigma = E[\xi_{\sigma,i}]$ and $\text{var}^Q[\xi_{\sigma,i}] = \dfrac{\frac{\mu_\sigma^2}{\nu}}{\left(1+\frac{\mu_\sigma}{\nu}\lambda_\sigma\right)^2} > \frac{\mu_\sigma^2}{\nu} = \text{var}[\xi_{\sigma,i}]$, i.e. under Q the agent expects jumps in σ_t^2 to have both a greater mean and variance. In conjunction with the increased frequency of jumps, this increase in mean and variance implies that, under Q, the agent expects more variation in σ_t^2 than under P.

We return to (30) and a comparison of the state dynamics under the risk-neutral and physical measures. Comparing the VARs under P and Q reveals two differences. The first is the change in the drift of y_t due to the term $-G_tG_t'\lambda_t$. This term also appeared in model 1, and is due to a shift in the mean of z_{t+1}, the gaussian innovations vector, when it is transformed under Q. This term changes the expected dynamics of y_t but does nothing to the distribution of innovations to y_t. Thus, it has no effect on time t conditional variances calculated under Q. The second difference in the VARs from P to Q is the change in the intensity and jump distributions of the compound Poisson processes, as delineated in (31)–(34). As noted, the change in the jump processes *does* increase the expected variation in x_t and σ_t^2 under Q. Comparing the system dynamics under P and Q clearly shows that the increased conditional variance in the jump processes J_{t+1}^Q should lead to a greater conditional variance of y_t under Q than under P. We are interested in the conditional variance of the market's capital gain, Δp_{t+1}, which was previously derived under P in (29). Using (31)–(34) it is easy to obtain an expression for $var_t^Q(\Delta p_{t+1})$:

$$\text{var}_t^Q(\Delta p_{t+1}) = \sigma_{p,t}^{2,Q} = \beta_{\sigma_p}^Q\sigma_t^2 + \beta_{\lambda_p}^Q\lambda_t + \varphi_d^2\,\tilde{\bar{\sigma}}^2(1-w_d) \tag{35}$$

where $\beta_{\sigma_p}^Q = \left(A_{x,m}^2\varphi_x^2 + A_{\sigma,m}^2\varphi_\sigma^2\right) + \varphi_d^2 w_d$ and $\beta_{\lambda_p}^Q = A_{x,m}^2(\sigma_x^2+\sigma_x^4\lambda_x^2)\exp\left(\frac{1}{2}\lambda_x^2\sigma_x^2\right)+A_{\sigma,m}^2\left(\dfrac{\mu_\sigma^2\nu}{(\nu+\mu_\sigma\lambda_\sigma)^2}+\left[\dfrac{\mu_\sigma\nu}{\nu+\mu_\sigma\lambda_\sigma}\right]^2\right)\left(1+\frac{\mu_\sigma}{\nu}\lambda_\sigma\right)^{-\nu}$. Though $\beta_{\sigma_p}^Q$ and $\beta_{\lambda_p}^Q$ appear complicated, they can be used to make two important points. First, notice that that $\beta_{\sigma_p}^Q$ is identical to β_{σ_p} from (29). This invariance results because this coefficient is a combination of loadings on *only* gaussian innovations and, as noted previously, the gaussian innovation conditional variances are invariant to the change of measure. The second point relates to $\beta^Q_{\lambda_p}$. It is not hard to see that when $\lambda_\sigma < 0$, i.e. the market price of variance risk is negative, we get $\beta^Q_{\lambda_p} > \beta_{\lambda_p}$, so that $\text{var}^Q_t(\Delta p_{t+1}) > \text{var}_t(\Delta p_{t+1})$. Thus, in the presence of jumps, a negative price of variance risk implies that the conditional variance of returns is greater under the risk-neutral probabilities.

4.2.1 Level Difference

Recall that in the discussion above we approximated the difference between summed variance under Q and P by the expression:

$$E^Q_t[\text{var}^Q_{t+1}(\Delta p_{t+2})] - E_t[\text{var}_{t+1}(\Delta p_{t+2})]$$

and defined the time t variance premium (vp_t) to be this quantity. We break up vp_t into two parts as follows:

$$\begin{aligned} vp_t &= \text{var}^Q_t(\Delta p_{t+1}) - \text{var}_t(\Delta p_{t+1}) + \\ &\quad \Big(E^Q_t[\text{var}^Q_{t+1}(\Delta p_{t+2})] - \text{var}^Q_t(\Delta p_{t+1})\Big) - \Big(E_t[\text{var}_{t+1}(\Delta p_{t+2})] - \text{var}_t(\Delta p_{t+1})\Big) \end{aligned} \tag{36}$$

The first line of this decomposition is the difference between the time t levels of conditional variance under Q and P. We refer to this difference as the "level difference". The second line is the difference between the drift of conditional variance under Q and P and we term this the "drift difference". Recall that when all innovations were gaussian, conditional variances were equivalent under Q and P. This meant that the variance premium resulted completely from the drift difference. As just shown, the presence of jumps induces also a level difference. This difference in conditional variances is very significant in our model calibration, constituting a majority of the variance premium. Moreover, the absence of the level difference in model 1 is a main reason behind its shortcomings in matching the data.

From (35) and (29) we obtain a simple expression for the level difference:

$$\text{var}^Q_t(\Delta p_{t+1}) - \text{var}_t(\Delta p_{t+1}) = \sigma^{2,Q}_{p,t} - \sigma^2_{p,t} = (\beta^Q_{\lambda_p} - \beta_{\lambda_p})\lambda_t \tag{37}$$

The expression shows that, so long as $\beta^Q_{\lambda_p} \neq \beta_{\lambda_p}$, the difference in conditional variances is simply a multiple of λ_t and perfectly "reveals" the jump intensity. As the level difference typically constitutes a majority of the total variance premium, it too will be driven largely by the jump intensity. This point is important, as it allows us to effectively 'observe' jump intensity with a high precision using option-derived quantities (i.e. the VIX), whereas inferring jump intensity from the return data alone is difficult. This is because jump intensity is latent and jumps are infrequent, so that even when the jump intensity is much higher than its mean, the occurrence of a jump is still far from certain. Thus, even if the jump intensity process is fairly persistent, the jump realizations need not be. The rarity of jump realizations and their lack of persistence makes the jump intensity process even more invisible. Thus, detecting and predicting the intensity process from return data alone can be tough. However, the jump intensity has a strong effect on the VIX and, as (37) shows, it comes through clearly in the level difference and in the variance premium. Thus, the use of option prices via the VIX makes the nearly invisible intensity process effectively observable. This is intuitive. Prices impute investors' fears of underlying risks going forward, and such fears will be strongly reflected in quantities such as the (one month) variance premium. In addition, as we show below in the section on predictability, the jump intensity *is* also reflected in the underlying market price, via the return risk premia. The affect of the jump intensity on the market risk premium accounts for the variance premium's predictive power for stock returns. This is an example of how derivative prices can be used in structural models to learn about the underlying asset's mean dynamics. Typical reduced-form option pricing models are silent on this since they either start directly in the risk-neutral world (where the mean return is the risk-free rate), or make assumptions regarding prices of risk in ways that may not impose the restrictions implied by preference-based models.

We now give an expression for the second line of (37), which is the drift difference, and then combine it with the level difference to get the whole variance premium.

4.2.2 Drift Difference

To get the drift difference, we substitute into the second line of (37) the expressions for conditional variance under P and Q given in (29) and (35). The two variance drifts are:

$$
\begin{aligned}
E_t[\mathrm{var}_{t+1}(\Delta p_{t+2})] - \mathrm{var}_t(\Delta p_{t+1}) &= \beta_{\sigma_p}\left[E_t(\sigma_{t+1}^2) - \sigma_t^2\right] + \beta_{\lambda_p}\left[E_t(\lambda_{t+1}) - \lambda_t\right] \\
E_t^Q[\mathrm{var}_{t+1}^Q(\Delta p_{t+2})] - \mathrm{var}_t^Q(\Delta p_{t+1}) &= \beta_{\sigma_p}^Q\left[E_t^Q(\sigma_{t+1}^2) - \sigma_t^2\right] + \beta_{\lambda_p}^Q\left[E_t^Q(\lambda_{t+1}) - \lambda_t\right]
\end{aligned}
$$

From this expression we can see that the drift difference is due to different expected dynamics in the state variables under P and Q. This was also the case under model 1. Risk-neutralizing the probabilities changes the agent's expectation of how the state variables evolve. The direction of this change depends on prices of risk. If the agent fears higher uncertainty/volatility, his risk-neutral expectation for σ_{t+1}^2 should be higher than his statistical (physical) expectation. Recall the dynamics of σ_t^2 under P:

$$E_t(\sigma_{t+1}^2) - \sigma_t^2 = \bar{\sigma}^2(1-\rho_\sigma) + (\rho_\sigma - 1)\sigma_t^2 + \mu_\sigma \lambda_t$$

The dynamics of σ_t^2 under Q are implicit in the VAR of (30), together with the properties of $J_{\sigma,t+1}$ under Q, (32) and (34):

$$\begin{aligned} E_t^Q(\sigma_{t+1}^2) - \sigma_t^2 &= \bar{\sigma}^2(1-\rho_\sigma) + (\rho_\sigma - 1)\sigma_t^2 - \varphi_\sigma^2 \lambda_\sigma \sigma_t^2 + E^Q[\xi_{\sigma,i}]\lambda_t^{Q,\sigma} \\ &= \bar{\sigma}^2(1-\rho_\sigma) + (\rho_\sigma - 1)\sigma_t^2 - \varphi_\sigma^2 \lambda_\sigma \sigma_t^2 + \frac{\mu_\sigma}{1 + \frac{\mu_\sigma}{\nu}\lambda_\sigma} \left(1 + \frac{\mu_\sigma}{\nu}\lambda_\sigma\right)^{-\nu} \lambda_t \end{aligned}$$

There are two changes. The term $\varphi_\sigma^2 \lambda_\sigma \sigma_t^2$ was observed also under model 1. It comes from the the gaussian innovations (from $-G_t G_t' \lambda$). The second change is due to the change in expected jump size and intensity. Both changes are directly related to λ_σ, the price of volatility risk. For both the gaussian and jump-induced changes, $\lambda_\sigma < 0$ implies an *increase* under Q in the drift in σ_t^2 as the discussion above suggested.

The final piece we need is:

$$\begin{aligned} E_t(\lambda_{t+1}) - \lambda_t &= \lambda_1 \left[E_t(\sigma_{t+1}^2) - \sigma_t^2\right] \\ E_t^Q(\lambda_{t+1}) - \lambda_t &= \lambda_1 \left[E_t^Q(\sigma_{t+1}^2) - \sigma_t^2\right] \end{aligned}$$

Putting the two pieces together, it is clear that the drift difference can be written as a function of σ_t^2 and λ_t:

$$\left(E_t^Q[\text{var}_{t+1}^Q(\Delta p_{t+2})] - \text{var}_t^Q(\Delta p_{t+1})\right) - \left(E_t[\text{var}_{t+1}(\Delta p_{t+2})] - \text{var}_t(\Delta p_{t+1})\right) = \beta_0^{dd} + \beta_\sigma^{dd}\sigma_t^2 + \beta_\lambda^{dd}\lambda_t$$

The expressions for β_0^{dd}, β_σ^{dd}, β_λ^{dd} are given in the appendix. Finally, we can get the whole variance premium by combining the drift difference and variance difference:

$$vp_t = \beta_0^{dd} + \beta_\sigma^{dd}\sigma_t^2 + \left(\beta_\lambda^{dd} + \beta_{\lambda_p}^Q - \beta_{\lambda_p}\right)\lambda_t \tag{38}$$

4.3 Risk Premia and Predictability

As discussed above, the (risk) factors driving the variance premium should also show up in return risk premia. It was noted that fear of a jump, in volatility or the long run component of cash flow growth, which has a strong influence on the (1 month) variance premium, should also show up in the risk premia demanded by investors. In particular, it should affect short horizon risk premia, especially if the jump intensity process is volatile but not very persistent. As such, the variance premium should predict excess stock returns, as was evidenced in the data discussion. We now address stock return predictability formally within the model.

To analyze predictability we need an expression for the conditional risk premium. The appendix shows that the conditional (log) risk premium can be written as:

$$E_t(r_{m,t+1}) - r_{f,t} = \beta_0^{re} + \beta_\sigma^{re}\sigma_t^2 + \beta_\lambda^{re}\lambda_t \tag{39}$$

where β_σ^{re} and β_λ^{re} are both positive when $\lambda_\sigma < 0$. (39) shows that σ_t^2 and λ_t are the drivers of conditional risk premia. Though λ_t is just a function of σ_t^2 in this model, we continue the practice of keeping its contribution separate so that we may focus on it in the presentation. We note that this means the same functional forms would continue to hold if λ_t were indeed a separate state variable[13]

Here we focus on the predictability of risk premia by the level difference, rather than the whole variance premium, since the resulting expressions are simpler and capture the same intuition. The appendix provides detailed derivations for predictability by both quantities. The projection on the level difference is:

$$r_{m,t+1} - r_{f,t} = \alpha + \beta_{\text{ld}}\left(\text{var}_t^Q(\Delta p_{t+1}) - \text{var}_t(\Delta p_{t+1})\right) + \epsilon_{t+1}$$

[13] For example, a minor and natural extension of the model could add an additional innovation to λ_t, i.e. let $\lambda_t = \lambda_0 + \lambda_1\sigma_t^2 + \varphi_\lambda z_{\lambda,t}$. This would make λ_t a separate state variable and reduce the (100%) correlation inside the model between σ_t^2, λ_t and vp_t.

It follows that

$$\begin{aligned}
\beta_{\mathrm{ld}} &= \frac{\mathrm{cov}\left(E_t(r_{m+1} - r_{f,t}) + \tilde{\epsilon}_{t+1},\ \mathrm{var}_t^Q(\Delta p_{t+1}) - \mathrm{var}_t(\Delta p_{t+1})\right)}{\mathrm{var}\left(\mathrm{var}_t^Q(\Delta p_{t+1}) - \mathrm{var}_t(\Delta p_{t+1})\right)} \\
&= \frac{\mathrm{cov}(\beta_0^{re} + \beta_\sigma^{re}\sigma_t^2 + \beta_\lambda^{re}\lambda_t,\ (\beta_{\lambda_p}^Q - \beta_{\lambda_p})\lambda_t)}{\mathrm{var}\left((\beta_{\lambda_p}^Q - \beta_{\lambda_p})\lambda_t\right)} \\
&= \frac{\beta_\sigma^{re} + \lambda_1\beta_\lambda^{re}}{(\beta_{\lambda_p}^Q - \beta_{\lambda_p})\lambda_1} \qquad (40)
\end{aligned}$$

where we use the fact that $\mathrm{cov}(\sigma_t^2, \lambda_t) = \lambda_1 \mathrm{var}(\sigma_t^2)$ and $\mathrm{var}(\lambda_t) = \lambda_1^2 \mathrm{var}(\sigma_t^2)$. Recall that when $\lambda_\sigma < 0$, we had $\beta_{\lambda_p}^Q - \beta_{\lambda_p} > 0$ and it is clear that $\beta_{\mathrm{ld}} > 0$. The intuition for this is simple. A high level of fear of jumps means investors demand a higher risk premia over the next period. The high expectation of jumps is clearly revealed by a high level difference, resulting in a positive projection coefficient.

The R^2 from the projection is simply $\beta_{\mathrm{ld}}^2 \frac{\mathrm{var}(\beta_{\lambda_p}^Q - \beta_{\lambda_p})}{\mathrm{var}(r_{m,t+1} - r_{f,t})}$. The appendix obtains expressions for these unconditional variances. The observation we wish to make from the expression for R^2, is that the variation in jump intensity accounts for the level difference's predictive power. Since the level difference accounts for a large portion of the variance premium, the variation in jump intensity is responsible for a large part of the variance premium's predictive power. This is a restatement of the discussion above regarding the influence of jump intensity on the risk premium.

5 Calibration Results

In calibrating the model we use the following guidelines. We assume a monthly decision interval. We would like to find a specification for the long run and various volatility shocks such that (i) once time-averaged to annual data the model's consumption and dividend growth statistics are consistent with salient features of the consumption and dividend dynamics (ii) the model generates consistent unconditional moments for asset prices such as the equity premium and the risk free rate (iii) the model's variance premium generates statistics as well as return projection results that are consistent with the data.

In Table V we provide the parameter specification for the model I economy of section 3. The model parameters are generally close to those in Bansal and Yaron (2004) and the volatility of q_t does not effect the annual cashflow dynamics in a significant manner. Table VI shows that the model captures quite well several key moments of annualized consumption and dividend growth. Table VII presents the model based asset pricing implications. For the market return and the risk free rate the reported statistics from the model are time averaged annual figures. The remaining statistics pertaining to the variance premium are given at the monthly frequency. The model generates a sizeable equity premium, and reasonable level of the risk free rate. The volatility of the market return is somewhat lower than the traditional full sample, but is consistent with the market volatility return given in Table I. It is interesting to note that the projection coefficients have the right sign but are somewhat too large relative to the data. Moreover, this baseline model has difficulties in generating the R^2 of 3-5% that is observed in data. In addition the level, and the variation of the variance premium are substantially lower than those observed in the data. In order to assess the sensitivity of this model to our specific parameterization we also considered the case of doubling the variation in q_t. However, even in this case, the variance premium in the model has difficulty matching the level and time variation in the data. As discussed earlier, the gaussian innovations in q_t ultimately lead to a gaussian variance premium and that dimension is difficult to reconcile with the data.

Table VIII specifies the parameter configuration for the model economy with jumps as described in section 4. Table IX provides the consumption and dividend growth moments. Again, the model parameters closely match those in Bansal and Yaron (2004) although we had to experiment with the jump parameters to settle on a configuration in which the volatility of consumption and dividends does not overshoot that in the data. A few words about the magnitude of the jumps and their frequency are in order. Under this calibration we chose the shape parameter of the jump distribution to be 1, thereby specializing the jumps to an exponential distribution. The size of the average jump in σ_t^2 is 1.75 times the unconditional mean of σ_t^2, which we normalize to a value of 1. The intensity of the jumps is given by $\lambda_1 = 0.75/12$. This directly translates into an average of 0.75/12 jumps per month or an average of 0.75 jumps per year. Thus, jumps are relatively infrequent, though they can be large when they occur.

Table X provides the relevant asset pricing results. The model generates a sizeable equity premium, and a reasonable level of the risk free rate. The volatilities of the market and the risk free rate are also broadly in line with those in the data. The interesting feature of the table is that now the model implied level, standard deviation, and skewness of the variance premium are much closer to that in the data. Moreover, the excess return predictability projection coefficients are now very close to those in the data and the corresponding R^2 is about 2.8%. In all the model does quite well in capturing the transient dynamics reflected in the variance premium. To better understand the quantitative workings of the model, Table X also displays the asset pricing results when the jump size is lowered to a half. It is easily seen that the level and predictability of the variance premium are now severely diminished, as is the equity premium. Finally, it is important to recognize that the preference parameters used here (e.g., risk aversion of ten) are similar in magnitude to those used and estimated successfully in other applications of the Long Run Risks model (e.g. Bansal, Kiku, and Yaron (2007)). This provides some cross-validation of these type of preferences.

6 Conclusion

This paper shows that the variance premium is useful for measuring agents' perceptions of uncertainty and for understanding how uncertainty is mapped by preferences into prices. We show that preference for early resolution of uncertainty and time variation in economic uncertainty are a minimal requirement for qualitatively generating a positive, time varying variance premium that predicts excess stock returns. We show that a Long Run Risks model with jumps in uncertainty can generate many of the quantitative features of the variance premium while remaining consistent with observed aggregate dynamics for dividends and consumption, as well as standard asset pricing data such as the equity premium and risk free rate.

One possible direction for generating interesting transient dynamics like the ones documented here is by generalizing the preferences to include features of ambiguity and desire for robustness. As Hansen and Sargent (2006) demonstrate a desire for robustness can lead to interesting time-varying misspecification risk premia components. The derivative related features of the data can thus be a fruitful ground for assessing the role these additional dimensions can provide in enhancing the model's ability to confront the data.

More generally, risk attitudes toward uncertainty play an important role in interpreting asset markets. The Long Run Risks has channels for several priced risk factors, including the level of uncertainty and its rate of change. An interesting direction for future research is determining the extent to which these risks are also important in the cross-section of returns. Ang, Hodrick, Xing, and Zhang (2006) is a recent work in this direction. Bansal, Kiku, and Yaron (2007) utilize an uncertainty factor in the cross-section of returns within the Long Run Risks framework, but are constrained to identify it based solely on cashflows. The evidence in this papers suggests that derivative markets and high frequency measures of variation should be very useful at identifying these risk factors. Interesting implications could therefore arise from jointly using cashflows and derivative markets to understand the influence of uncertainty on the cross-section.

7 Appendix

TBA

References

Ang, Andrew, Robert Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The Cross-Section of Volatility and Expected Returns, *Journal of Finance* 51, 1, 259–299.

Bansal, Ravi, Varoujan Khatchatrian, and Amir Yaron, 2005, Interpretable asset markets?, *European Economic Review* 49, 531–560.

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.

Bloom, Nick, 2007, The Impact of Uncertainty Shocks, Working paper, Stanford University.

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.).

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

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*. (Princeton University Press Princeton, New Jersey).

Campbell, John Y., and Samuel B. Thompson, 2007, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, *Review of Financial Studies* forthcoming.

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

Cochrane, John H., 1999, Portfolio advice for a multifactor world, *Economic Perspectives* XXXIII(3), (Federal Reserve Bank of Chicago).

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

Duffie, D., J. Pan, and K. J. Singleton, 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica* 68, 1343–1376.

Epstein, Larry G., and 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, 2007, Affine General Equilibrium Models, *Management Science* forthcoming.

Eraker, Bjorn, and Ivan Shaliastovich, 2007, An Equilibrium Guide to Designing Affine Pricing Models, *Mathematical Finance* forthcoming.

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

Hansen, Lars, and Thomas Sargent, 2006, Fragile Beliefs and the Price of Model Uncertainty, Working paper, .

Heston, Steven L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies* 6, 2, 327–343.

Jiang, George, and Yisong Tian, 2005, Model-Free Implied Volatility and Its Information Content, *Review of Financial Studies* 18, 1305–1342.

Lettau, Martin, Sydney Ludvigson, and Jessica Wachter, 2007, The Declining Equity Premium: What Role Does Macroeconomic Risk Play?, *Review of Financial Studies* Forthcoming.

Stambaugh, Robert F., 1999, Predictive Regressions, *Journal of Financial Economics* 54, 375–421.

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

Table I

Summary Statistics

	Excess Returns		Variances			
	S&P 500	NYSE-AMEX-NASDAQ	VIX^2	Fut^2	Ind^2	$Daily^2$
Mean	0.528%	0.526%	33.30	22.17	14.74	20.69
Median	0.957%	1.023%	25.14	14.19	8.99	13.51
Std.-Dev.	4.01%	4.13%	24.13	22.44	15.30	21.95
Skewness	-0.635	-0.836	2.00	2.62	2.78	2.68
Kurtosis	4.217	4.547	8.89	11.10	13.26	11.91
AR(1)	-0.04	0.02	0.79	0.65	0.73	0.62

Table I presents descriptive statistics for excess returns and realized variances. The sample is monthly and covers 1990m1 to 2007m3.

Table II

Conditional Volatility

Dept. Variable	Regressors					
	X1	X2	intercept	β_1	β_2	R^2
Daily^2_{t+1}	Daily^2_t	MA(1)	3.694	0.820	-0.350	0.40
(t-stat)			(2.529)	(13.540)	(-3.536)	
Ind^2_{t+1}	Ind^2_t	VIX^2_t	0.098	0.400	0.261	0.59
(t-stat)			(1.165)	(5.534)	(5.689)	
Fut^2_{t+1}	Ind^2_t	VIX^2_t	-0.889	0.560	0.290	0.59
(t-stat)			(-0.51)	(2.71)	(8.27)	

Table II presents regressions of realized variance measures on lagged predictors. The forecast from the third regression is used as the conditional realized variance estimate.The sample is monthly and covers 1990m1 to 2007m3.

Table III

Properties of the Variance Premium

	VP(BZ)	VP(Ind–forecast)	VP(Daily–MA(1))	VP(Fut–forecast)
Mean	18.41	18.50	12.52	11.20
Median	14.17	14.95	7.92	8.81
Std.-Dev.	15.34	13.76	14.37	7.64
Minimum	-26.05	3.88	-4.02	3.27
Skewness	2.17	2.04	2.50	2.43
Kurtosis	12.06	9.26	12.88	12.16
AR(1)	0.50	0.69	0.54	0.65

Table III presents statistics for various measures of the variance premium. The sample is monthly and covers 1990m1 to 2007m3.

Table IV

Return Predictability by the Variance Premium

Regressors					
X1	X2	intercept	β_1	β_2	adj. $R^2(\%)$
VP_t		-1.22	0.73		0.90
(t-stat)		(-.28)	(2.12)		
VP_{t-1}		-6.918	1.236		3.44
(t-stat)		(-1.29)	(3.82)		
VP_t	$\log{(\text{P/D}^*)}_t$	94.41	1.10	-41.62	5.35
(t-stat)		(3.82)	(2.81)	(-3.76)	
VP_{t-1}	$\log{(\text{P/D}^*)}_t$	94.15	1.59	-43.81	8.58
(t-stat)		(3.95)	(4.95)	(-4.22)	
VP_t	$\log{(\text{P/E})}_t$	144.14	1.38	-49.10	7.44
(t-stat)		(3.14)	(3.00)	(-3.11)	
VP_{t-1}	$\log{(\text{P/E})}_t$	165.69	2.08	-58.54	12.55
(t-stat)		(3.49)	(4.84)	(-3.58)	

Table IV presents return predictability regressions. Reported t-statistics are Newey-West corrected with four lags. P/D is the price-dividend ratio The The sample is monthly and covers 1990m1 to 2007m3.

Table V

Calibration Model I: Parameter Configuration

Preferences	δ	γ	ψ		
	0.999	10	1.5		
Consumption	μ_c	ρ_x	φ_x	φ_c	
	0.0015	0.977	3.12e-4	0.00546	
Dividends	μ_d	ϕ	φ_d		
	0.0015	3	2.46e-2		
Volatility	$\bar{\sigma}^2$	ρ_σ	$\bar{q}$	ρ_q	φ_q
	1	0.93	0.211	0.7	0.632

Table V presents the parameter configuration used to calibrate the model I of section 3.

Table VI
Consumption & Dividend Growth Dynamics

Statistic	**Data**		**Model**
$E[\Delta c]$	1.96	(0.34)	1.89
$\sigma(\Delta c)$	2.21	(0.38)	2.53
$AC(1)$	0.44	(0.13)	0.51
$E[\Delta d]$	0.74	(1.18)	1.86
$\sigma(\Delta d)$	11.0	(1.92)	9.76

Table VI presents several moments for consumption and dividend growth at the annual frequency. The data corresponds to the period from 1930 to 2002 and the standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags. The model implied statistics are based on simulating a model at the monthly frequency and time averaging the statistics to annual figures.

Table VII

Calibration Results: Model I

Statistic	**Data**		**Model**
$E[R_{m,t+1} - R_{f,t}]$	7.51	(2.10)	7.44
$E[R_{f,t}]$	0.76	(0.27)	1.12
$\sigma(R_{m,t})$	20.1	(1.88)	21.5
$\sigma(R_{f,t})$	1.12	(0.22)	1.68
$\sigma(\mathrm{var_t}(\Delta p))$	17.15		18.99
$AC(\mathrm{var_t}(\Delta p))$	0.81		0.88
$E[VP_t]$	11.20		0.49
$median[VP_t]$	8.81		0.33
$\sigma(VP_t)$	7.64		1.59
$skew(VP_t)$	2.43		0
$AC(1)$	0.65		0.70
$\beta(VP_{t-1})$	1.24		2.77
$\bar{R}^2$	3.44		0.83

Table VII presents several asset pricing moments from calibrating the model in section 3. The data for market and risk free rate corresponds to the period from 1930 to 2002. The data for the variance premium reproduces the statistics from Table III. Standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags. The model implied statistics for the market and the risk free rate are based on simulating a model at the monthly frequency and time averaging the statistics to annual figures. The statistics for the variance premium are at the monthly frequency.

Table VIII

Calibration Model II: Parameter Configuration

Preferences	δ	γ	ψ		
	0.999	10	1.5		
Consumption	μ_c	ρ_x	φ_x	φ_c	w_c
	0.0015	0.975	3.12e-4	0.00664	0.8
Dividends	μ_d	ϕ	φ_d	w_d	
	0.0015	3	0.0299	0.5	
Volatility	$\tilde{\bar{\sigma}}^2$	$\tilde{\rho}_\sigma$	φ_σ		
	1	0.90	0.04		
Jumps	λ_0	λ_1	ν	μ_σ	σ_x^2
	0	0.75/12	1	1.75	4.68e-4

Table VIII presents the parameter configuration for the model with jumps as described in section 4.

Table IX

Consumption & Dividend Growth Dynamics

Statistic	**Data**		**Model**	$\frac{1}{2}$ **jump**
$E[\Delta c]$	1.96	(0.34)	1.95	1.88
$\sigma(\Delta c)$	2.21	(0.38)	2.49	2.45
$AC(1)$	0.44	(0.13)	0.49	0.49
$E[\Delta d]$	0.74	(1.18)	2.04	1.77
$\sigma(\Delta d)$	11.0	(1.92)	10.48	10.48
$AC(1)$	0.21	(0.13)	0.38	0.38
$corr(\Delta c, \Delta d)$	0.60	(0.14)	0.31	0.31

Table IX presents several moments for consumption and dividend growth at the annual frequency. The data corresponds to the period from 1930 to 2002 and the standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags. The model implied statistics are based on simulating a model at the monthly frequency and time averaging the statistics to annual figures.

Table X
Calibration Results: Model II

Statistic	**Data**		**Model**	$\frac{1}{2}$ **jump**
$E[R_{m,t+1} - R_{f,t}]$	7.51	(2.10)	7.61	4.67
$E[R_{f,t}]$	0.76	(0.27)	1.51	1.67
$\sigma(R_{m,t})$	20.1	(1.88)	18.8	15.7
$\sigma(R_{f,t})$	1.12	(0.22)	1.55	1.18
$\sigma(\mathrm{var_t}(\Delta p))$	17.15		27.09	8.78
$AC(\mathrm{var_t}(\Delta p))$	0.81		0.91	0.89
$E[VP_t]$	11.20		9.04	0.36
$median[VP_t]$	8.81		4.96	0.27
$\sigma(VP_t)$	7.64		12.49	0.23
$skew(VP_t)$	2.43		5.64	3.50
$AC(1)$	0.65		0.91	0.89
$\beta(VP_{t-1})$	1.24		0.79	10.15
$\bar{R}^2$	3.44		2.80	0.22

Table X presents several asset pricing moments from calibrating the model in section 4. The data for the market and risk free rate corresponds to the period from 1930 to 2002. The data for the variance premium reproduces the statistics from Table III. Standard errors are calculated using the Newey-West variance-covariance estimator with 4 lags. The model implied statistics for the market return and the risk free rate are based on simulating a model at the monthly frequency and time averaging the statistics to annual figures. The statistics for the variance premium are at the monthly frequency.