Real-Time Learning, Macroeconomic Uncertainty, and the Variance Risk Premium*

Daniele Bianchi†

Abstract

The variance risk premium represents the compensation paid to index option sellers for the risk of losses following upward movements in realized market return volatility. Common wisdom connects these spikes with elevated uncertainty on economic fundamentals. I incorporate this link within a single-agent general equilibrium model, embedding real-time learning on state variables and parameters. I show that infrequent, large and relatively transitory macroeconomic uncertainty shocks produce a sizable and volatile variance risk premium. These shocks coincide with major events such as the LTCM/Russian crisis, the onset of the second Gulf War, and the great financial crisis of 2008-2009. I compute macroeconomic uncertainty as the dispersion of the agent’s belief about the expected growth rate of consumption. Its time-varying nature reflects in the variance risk premium, generating short-term predictability for market excess returns, consistent with the data. In addition, the model matches the higher order moments of the realized equity premium, with a reasonably low level of relative risk aversion equal to five. I finally provide evidence that parameter uncertainty may represent an extra-source of risk which is priced in equilibrium. In fact, a model with parameter learning and standard CRRA preferences, matches around half of the historical variance risk premium.

Keywords: Variance Risk Premium, Real-Time Learning, Parameter Uncertainty, Recursive Preferences

JEL codes: G12, G13, E44, C11

*This version: 11 November 2014. For the latest version see my webpage. I would especially like to thank my committee, Carlos M. Carvalho, Massimo Guidolin and Paolo Colla, for extremely valuable suggestions and support. I also thank Luc Bauwens (discussant), Fabio Canova, Chris Sims, Francesco Ravazzolo, Roberto Casarin and participants at the ESOFBE Meeting, Oslo, 2013. I also benefited from conversations with Omar Barbiero, Mascia Bedendo, Daniele Bieber, Pierre Collin-Dufresne, Robert Dittmar, Andres Donangelo, Travis L. Johnson, Giovanni W. Puopolo, Jan Schneider, Claudio Tebaldi, and Haoxi Yang. The paper is based on the first chapter of my PhD dissertation at the department of Finance, Bocconi University.

†University of Warwick, Finance Group, Coventry, CV32 74L, UK daniele.bianchi@wbs.ac.uk
1 Introduction

Macroeconomic uncertainty and periods of extreme turmoil have recently plagued financial markets. This has prompted an increasing interest in equity index options aimed at hedging the risk of losses following upward movements in market return volatility.\(^1\) The compensation paid by investors for such options – the so-called variance risk premium – is increasing in the implied variance of market returns. The latter, is commonly referred to as the “fear gauge” for the investors’ uncertainty on economic fundamentals. Figure 1 displays the variance risk premium against several survey-based indexes of macroeconomic uncertainty.\(^2\) The higher the uncertainty about the economy, the larger the variance premium. Vector autoregression (VAR) estimations also suggest that shocks in macroeconomic uncertainty have a considerable impact, generating a rise and fall in the variance risk premium over the following months. In this paper, I incorporate this link within a single-agent general equilibrium model that features real-time learning about the parameters of economic fundamentals. As such, the investor is burdened with the same problems faced by the econometrician and must simultaneously learn over time about state variables and parameters. The model matches the time series dynamics and the unconditional moments of the premium enclosed in market index options. At the same time, with a reasonably low level of relative risk aversion equal to five, the model also matches higher order moments of the realized equity premium and salient properties of cash-flows, consistent with the data.

The presence of a large variance risk premium has been extensively reported in the literature both in the time series and in the cross section. Bakshi and Kapadia (2003) provide evidence of a substantial market variance premium from a set of S&P500 options. Bakshi and Madan (2006) show that this premium may be linked to investors’ risk aversion and the higher moments of market return. Carr and Wu (2009) demonstrate that the index options premium is on average sizable across different stocks and indexes, and can not be explained by standard risk factors such as market, size, value and momentum. Bali and Hovakimian (2009) show that there is a significant and positive relation between the variance premium and the cross-section of expected stock returns. Similarly, Han and Zhou (2013) provide evidence that a long-short strategy built sorting stocks based on their variance risk premia generates a positive and statistically significant excess return

\(^1\)The risk of losses comes from the fact that jumps in volatility are seen as unfavorable shocks to the investment opportunity set, reducing for instance the optimal Sharpe ratio for a given expected return.

\(^2\)Throughout the paper I use the terms macroeconomic uncertainty, economic uncertainty and economy-wide uncertainty, interchangeably. I also use structural learning, structural uncertainty, parameter uncertainty and parameter learning as synonym.
which can not be explained by standard risk factors and firm-specific characteristics. Although the existence of a large variance premium is relatively well-established, much less is known about its link with macroeconomic fundamentals. This is the focus of my work.

In this paper, I show that shocks in macroeconomic uncertainty – defined as the dispersion of the agent’s belief on the expected growth rate of consumption – generate a large and volatile variance risk premium. These spikes in uncertainty are large but relatively infrequent, transitory, and occur at the end of the 1990s (LTCM/Russian crisis), the period 2001-2002, (9/11, Worldcom, Enron, dot.com bubble and the onset of the second Gulf War), and across the financial crisis of 2008-2009. I also show that the variance premium is increasing in parameter uncertainty. In fact, structural uncertainty represents a non-diversifiable source of risk which is manifested through a more pessimistic pricing kernel, namely increasing the discount rates. This feature, endogenously makes the price-dividend ratio more reactive in “bad” times rather than in “good” times, inflating the counter-cyclical behavior of equity returns volatility, as well as increasing the subjective probability of a positive shock in market return variance. These two effects together, enlarge the variance risk premium in equilibrium. A further implication of the model is that the index options premium reflects the time-varying nature of macroeconomic uncertainty shocks, which are counter-cyclical and negatively in-sample correlated with the realized equity premium. This negative relation reverses out-of-sample, generating short-term market excess returns predictability as we find in the data and consistent with previous evidence such as Bali and Hovakimian (2009) and Bollerslev, Tauchen, and Zhou (2009).

The framework designed in this paper includes recursive Kreps-Porteus preferences and real-time learning on state variables and structural parameters. The latent states are the conditional expected growth rate of consumption and the regime of macroeconomic uncertainty driving the dynamics of the variance risk premium. The structural parameters are the unconditional expected growth rate of cash-flows, their idiosyncratic risk, the leverage factor on dividend growth and the transition probabilities of the Markov-switching uncertainty regime. Such a rich learning dynamics directly influences the variance premium along two main directions. First, parameter learning induces low frequency fluctuations in the conditional distribution of consumption growth, generating an endogenous long-run type of risk (see Collin-Dufresne, Johannes, and Lochstoer 2013 for

---

3Simply put, the agent puts more weight, with respect to the full-information case, on the low continuation values of recursive utility during periods of high macroeconomic uncertainty.


5These parameters are those the reference literature assume to be observable by the agent.
a complete discussion). This endogenous persistence inflates the duration of uncertainty shocks as a result of their effect on the signal-to-noise ratio. The mechanism is as follows. A positive shock in macroeconomic uncertainty makes the agent more concerned about current consumption, affecting her posterior beliefs on the structural parameters. These beliefs revision generate a permanent shock in the conditional dynamics of the expected consumption growth. As such, highly transitory spike in uncertainty turn out to have a persistent effect because of their impact on the agent beliefs. Second, parameter learning alters the agent’s belief on a high uncertainty regime. To illustrate, suppose the agent realizes that the transition probability from a state of low to high macroeconomic uncertainty is higher than expected. This enlarges the perceived probability of the latter, then increases the variance risk premium in equilibrium.

In addition to salient moments of the variance risk premium and its short-term predictive power for excess returns, the model also matches a broad set of features of the market excess returns such as higher order moments and the counter-cyclical conditional variance. These characteristics have direct implications for index option prices and premia (see Bakshi and Madan 2006 for more details). The model also replicates a high and volatile market excess return, a low real risk-free rate, and a pro-cyclical persistent log price-dividend ratio, with a reasonably low level of relative risk aversion of five. Counter-cyclical return variance and negative skewness comes as a by-product of the time-varying flow of information inherited from parameter learning. Indeed, in bad states consumption growth becomes more informative on the uncertainty regime. As such, the price-dividend ratio becomes more responsive to belief updates in bad times, leading to both counter-cyclical volatility of market returns, and learning asymmetries. The latter generates a negative skew in the unconditional agent’s belief which reflects in a pro-cyclical and negatively skewed market excess return. The model-implied variance risk premium turns out to be stable across different levels of risk aversion.

Finally, I provide evidences that parameter uncertainty may effectively represent an extra source of non-diversifiable risk which increases the premium to hedge for the variance risk. In fact, even with time-separable CRRA preferences, the model with parameter learning still matches around half of the variance risk premium. The specific role of structural learning is tested by comparing the model results with a rational expectations benchmark. At a more general level, these results, may help reach a better understanding of the economics behind the large and volatile premia for both the price and the

---

6 A rational expectations benchmark is a calibrated version of the model in which structural parameters are set to be equal to end-of-sample estimates computed from the model with parameter learning. This is equivalent to calibrating the model parameters to those values computed by the agent once the whole historical information is made available.
variance risk, namely the equity and the variance risk premia.

This paper fits into a growing literature that aims to solve the so-called variance risk premium puzzle, namely the failure of standard consumption-based asset pricing models to explain the large and volatile historical index options premium. Indeed, representative agent models involving time-separable preferences imply a null variance premium. Likewise, both a classic long-run risk model with stochastic volatility and the habit-formation of Campbell and Cochrane (1999) rule out price variance risk and the corresponding hedging demand. Most of the existing literature typically builds on reduced form option pricing models such as Brodie, Chernov, and Johannes (2007), Todorov (2010), Bollerslev, Gibson, and Zhou (2011) and Bollerslev and Todorov (2011), just to cite a few. Brodie et al. (2007) show that jumps in market prices and volatility bring an economically relevant risk which is significantly priced in options. Todorov (2010) suggests that, in fact, the variance risk premium may be a direct consequence of two properties of market indexes such as jumps in prices and stochastic volatility. Similarly, Bollerslev and Todorov (2011) show that a large fraction of the variance risk premium involves a compensation for rare events measured as medium sized jumps in high-frequency returns. Bollerslev et al. (2011) demonstrate that the variance risk premium helps predicting future stock market returns and relates to a set of macro-finance variables that reflect a business cycle effect. Similarly, Corradi, Distaso, and Mele (2013), reveal that the variance risk premium is strongly negatively related to the business cycle. The equilibrium model in this paper incorporate these characteristics and is consistent with these findings.

Despite this large amount of empirical evidence, few papers link macroeconomic fundamentals and preferences with index option prices in equilibrium-based settings. These include Liu, Pan, and Wang (2005), Bekaert, Engstrom, and Xing (2009), Bollerslev et al. (2009), Benzoni, Collin-Dufresne, and Goldstein (2011), Drechsler and Yaron (2011), Miao, Bin, and Zhou (2012), and Drechsler (2013). Bollerslev et al. (2009) find that time-varying volatility of volatility in consumption growth provides an explanation of the short-term predictive power of the variance risk premium. By using an extended long-run risk model, Drechsler and Yaron (2011) show that non-Gaussian shocks in both the conditional expected growth rate of consumption and macroeconomic risk may help explain some of the key unconditional moments of the variance premium. Benzoni et al. (2011) apply a similar setting to explain the inconsistency between the smooth dynamics of economic fundamentals and the jump-like behavior of stock market returns. Bekaert et al. (2009) fit a long-run risk dynamic within an external habit formation setting, showing that time-varying volatility in economic fundamentals generates a sizable market return variance. Liu et al. (2005) show that model uncertainty about rare events may explain some
characteristics of options such as implied volatility skewness. Drechsler (2013) provides evidence that time-varying concerns on model uncertainty may generate a set of key features of index options premium. Similarly, Miao et al. (2012) use a single-agent endowment economy with smooth ambiguity preferences and incomplete information on the business cycle, showing that a high ambiguity aversion may help explain unconditional moments of the variance risk premium.

These studies rely on two main common features. First the agent can observe the structural parameters governing the dynamics of economic fundamentals. Second, there is a persistent component in the dynamics of the conditional expected growth rate of consumption and/or macroeconomic risk. However, the existence of such persistent components is still under debate (see Hansen 2007, Hansen, Heaton, and Li 2008, Hansen and Sargent 2010, and Sargent 2007 among the others). More generally, given the forward looking nature of equilibrium conditions, unless the agent directly observes over time the low-frequency component driving consumption, it is not clear how persistent shocks may be reflected on equilibrium prices. Bekaert and Hoerova (2013), for instance, cast some doubt on the effective persistence in the dynamics of macroeconomic risk. Figure 2 confirms this low-persistence result showing the squared residuals from an AR(12) fitted on consumption growth. Appealing to the standard idea that the econometrician has a smaller information set than the econometrician is also a little too restrictive in this context. In the long-run risk-based models, indeed, the exact conditioning information of the agents is needed in order to derive the asset pricing implications of the model.

In this paper, I differ from existing works along two key dimensions. First, the investor fully acknowledges uncertainty on the model parameters and learns about them in real-time. In equilibrium, this learning scheme endogenously creates a subjective long-run type of risk which is reflected on both equity and variance risk premium. This bridges the gap between ex-ante transitory shocks and their ex-post persistent effect on the dynamics of consumption growth. Second, I do not impose a priori any particular level of persistence in the dynamics of state variables. Both parameter estimates and model testing show that, in fact, consumption growth may be closely perceived as a conditional i.i.d. process for around half of the sample. This makes consumption far less predictable, consistent with the argument in Campbell and Beeler (2012). Without having to rely on exogenously imposed persistence, the model is still able to generate a sizable variance risk premium and replicate the short-term predictive power for market excess returns, and

---

Footnote: In particular Hansen et al. (2008) point out that: “Many of the statistical challenges that plague econometricians presumably also plague market participants. Naive application of rational expectations equilibrium concepts may endow investors with too much knowledge about future growth prospects. Learning and model uncertainty are likely to be particular germane to understanding long-run risk.”
at the same time matching higher order moments of equity returns. This endogenous persistence also increases the duration of shocks in the level of economic uncertainty which are transitory per-se, having however a relatively long-lasting effect. Above and beyond getting rid of the likely too restrictive assumptions the reference literature is based on, the model is able to match the finite sample historical dynamics of the variance risk premium. In particular, the model turns out to fairly closely match the extreme values characterize the conditional dynamics of the variance risk premium. As far as the reference literature is concerned, the dynamic feature has not been closely replicated within a general equilibrium setting.

The paper also contributes to a recent literature aimed at explaining other asset pricing puzzles as a result of structural uncertainty. These include Lewellen and Shanken (2002), Weitzman (2007), Bakshi and Skoulakis (2010), Johannes, Lochstoer, and Mou (2011), and Collin-Dufresne et al. (2013). These works consider the equity risk premium in isolation. However, as pointed out in Bollerslev and Todorov (2011), there is a tight connection between the equity and the variance risk premia as they tend to co-move. Thus, any satisfactory equilibrium asset pricing model must be able to generate both premia. I fill this gap by showing that parameter uncertainty may convey enough information to jointly help explain a relevant fraction of both risk premia.

Methodologically, this paper represents one of the first attempts, together with Johannes et al. (2011) and Collin-Dufresne et al. (2013), to apply simultaneous Bayesian learning about state variables and parameters to general equilibrium asset pricing models. This contrasts with most existing works that focus on learning latent states or parameters, alternatively. Simultaneous learning about multiple unknowns is able to better capture uncertainty on the dynamics of consumption-based asset pricing models. In this paper, for instance, both the size, the magnitude and the timing of uncertainty shocks are estimated and not calibrated a priori, making the model more flexible and robust, as I let the data directly speak on the nature of uncertainty shocks. As shown in Martin (2013) and Chen, Dou, and Kogan (2013), to fully account for parameter uncertainty may represent a crucial advantage. In fact, assuming complete knowledge of the parameters induces fragility in full-information rational expectations as small changes in the input parameters, such as those of jump-like shocks, may lead to sensibly different outputs. Finally, parameter learning endogenously generates non-normalities in the con-

---

ditional dynamics of consumption growth, still preserving the standard assumption of conditionally Gaussian innovations in the data generating process.

The rest of the paper proceeds as follows. Section 2 introduces the variance risk premium, its measurement and the data, as well as preliminary empirical evidence on the relationship between macroeconomic uncertainty and the variance risk premium. Section 3 presents the model. Section 4 reports the empirical results, the parameters estimates and briefly discuss the pricing mechanism underlying the model. Section 5 concludes.

2 Data and Definitions

The variance risk premium is defined as the difference between the risk-neutral and the physical expectations of aggregate stock market return variance for a given horizon \( \tau \);

\[
VRP_{t,t+\tau} = E_t^Q [RV_{t,t+\tau}] - E_t^P [RV_{t,t+\tau}]
\]

in which \( RV_{t,t+\tau} = \tau^{-1} \int_t^{t+\tau} \sigma_m^2(s) ds \) with \( \sigma_m^2(s) \) the local variance at time \( s \), \( E_t^P [\cdot] \) and \( E_t^Q [\cdot] \) the conditional expectation operator at time \( t \) under the physical and the risk-neutral measure, respectively. Carr and Wu (2009), Britten-Jones and Neuberger (2000) and Demeterfi, Derman, Kamal, and Zou (1999), show that the risk-neutral expectation can be approximated as the variance swap rate \( SW_{t,t+\tau} \) defined over the period \([t, t+\tau]\), which can be computed as the value of a static portfolio of options

\[
E_t^Q [RV_{t,t+\tau}] = SW_{t,t+\tau} = \frac{2}{\tau B_t (t+\tau)} \left( \int_0^{F_{t,t+\tau}} \frac{1}{K^2} P_{t,t+\tau}(K) dK + \int_{F_{t,t+\tau}}^{\infty} \frac{1}{K^2} C_{t,t+\tau}(K) dK \right)
\]

where \( F_{t,t+\tau} \) the futures price, \( B_t (t+\tau) \) the time-t price of a bond paying one dollar at time \( t+\tau \), \( P_{t,t+\tau}(K) \) and \( C_{t,t+\tau}(K) \) respectively the price of an out-of-the-money put and call option at time \( t \) with maturity \( t+\tau \) and strike price \( K \). This approximation is referred to as the implied variance of aggregate market returns. Theoretically \( SW_{t,t+\tau} \) relies on an increasing number of options with strikes spanning from zero to infinity. In practice, the variance swap rate is proxied by the CBOE implied volatility (or VIX), which is based on a finite number of high-liquid S&P500 out-of-the-money index options. I focus on a one-month variance premium such that \( \tau = 1 \). Since the VIX is reported as an annualized volatility measure, I square and divide it by 12 to return it in variance terms. Despite the fact that the VIX index is subject to approximation error, it represents the standard in the financial industry.
The physical expectation relies on an empirical measure of the aggregate stock market variance. To construct such a measure, I use high-frequency log returns, summing up 78 intra-day five-minute squared returns covering a standard trading day from 9:30 am to 4:00 pm. Data are obtained from TICKDATA. The sample period is 1990:01 - 2013:01, monthly. Note that, high-frequency intra-day returns lead to a much more precise approximation of the true (unobserved) return variance than more traditional daily market returns (Meddahi 2002 and Barndorff-Nielsen and Shepard 2002).\footnote{The main limitation of the sample length comes from the fact that the VIX index is available since 1990:01. The starting point 1990:01 refers to the new method to compute the VIX. See http://www.cboe.com/micro/VIX/vixintro.aspx for more details. Note that, market micro-structure issues such as bid-ask spreads, non-synchronous trading effects, as well as price discreteness, may limit the usefulness of very high frequency prices. Bollerslev et al. (2009) point out that micro-structure noise implies that the underlying semi-martingale assumption for aggregate returns is violated at the very highest sampling frequencies.}

Although a high-frequency discretized measure of the realized variance is widely accepted in the literature, the method for constructing the physical expectation \( E_t^p [RV_{t,t+1}] \) is not unique. Following Drechsler and Yaron (2011) and Drechsler (2013) I use as conditional expectation a simple linear projection of current realized variance \( RV_{t,t+1} \), on both implied and realized lagged returns variance (see the appendix).\footnote{For the sake of completeness I preliminary investigate other alternative measures as in Bollerslev et al. (2009) and Han and Zhou (2013) (see the appendix for more details).} The one-step ahead forecast from this linear regression, approximates the conditional expectation of the stock market variance under the physical measure. This projection-based method gives a positive variance premium for the majority of the sample. In fact, a completely model-free approach (meaning the conditional expectation be equal to the lagged realized variance \( RV_{t-1,t} \)) generates a large negative value especially across the great financial crisis of 2008/2009. Since the variance risk premium represent the insurance compensation an investor is willing to pay to hedge for positive shocks in market volatility, a negative variance premium may not be easily interpretable.\footnote{A possible explanation may be related to a structural imbalance between supply and demand of options during that period. In fact, a negative premium means that the agent would realize a positive return on a long-short strategy on volatility across 2008/2009. This paper does not add anything to this debate.}

Market returns correspond to the value-weighted return of NYSE,AMEX and NASDAQ, returned in real terms by using the CPI deflator. Data are from CRSP and the FREDII database held by the St. Louis Fed. Per-capita consumption is obtained summing up consumer expenditures on non-durable and services, adjusted for the population, and made it in real terms. Data are from the NIPA database. Growth rates are constructed by taking first differences of the corresponding log series. The aggregate dividends are computed as in Campbell and Beeler (2012) and corrected for repurchases.
following Bansal, Dittmar, and Lundblad (2005). The growth rate of aggregate dividends is transformed in real terms by subtracting the log inflation, which is obtained from the CPI deflator. Nominal yields to calculate the risk-free rates are from Ibbotson as the 30 days T-Bill return. The forward looking perspective of the model requires the use of an ex-ante measure of the risk free rate. This is approximated by projecting the one-step ahead real T-Bill yield on past log inflation and current nominal interest yield. I leave more details on the data to the appendix.

Table 1 provides a set of descriptive statistics for both the variance measures, total stock market returns, the expected realized variance under $\mathbb{P}$, and the corresponding variance risk premium. As we would expect, both the implied and the realized market variances sensibly increase, on average, across the period 2008/2009. Furthermore, $RV_{t,t+1}$ doubles its volatility by including the last part of the sample. This is also true for the implied variance $IV_{t,t+1}$. Putting together these effects, the average premium hikes from 15.82 to 18.36, and the unconditional volatility arises from 10.76 to 12.69, monthly. However, despite the negative jump that occurred across the recent financial crisis, both the skewness and excess kurtosis of the variance premium are comparable across the pre-crisis (1990:01-2007:12) and the full sample. The fact that the premium is positive on average, is consistent with the idea that options on index-returns volatility represent insurance against sudden upward movements in market variance.\textsuperscript{12} As the table shows, the variance risk premium is not normally distributed, in fact, a standard Jarque-Bera test strongly rejects the null of $VRP_{t,t+1}$ be Gaussian distributed.\textsuperscript{12}

Table 2 reports a set of excess return predictive regressions. The dependent variable is the $k = 1, ..., 24$ steps ahead aggregate market excess returns, and the independent is the current variance risk premium. Monthly returns are overlapping. I report two sets of Bayesian regressions. Top panel shows the results from a standard normal-inverse-gamma conjugate Bayesian regression with Gaussian distributed error terms. Bottom panel shows the results of a robust version with $t$-distributed innovations, which allows to both consider the impact of outliers and acknowledge the uncertain nature of parameters estimates.\textsuperscript{13} The table reports posterior median estimates, as well as the confidence

\textsuperscript{12}These jumps in volatility are seen as negative shocks in the investment opportunity set. As such, risk-averse investors are even willing to loose money to insure such disliked events.

\textsuperscript{13}More details on the Bayesian regressions can be found in a separate online appendix.
intervals at the 95% level for each of the slope parameters.

The variance premium shows a significant predictive power at the short-term with a median adjusted $R^2$ of almost 7% at the quarterly horizon. Consistently with Bollerslev et al. (2009) the forecasting ability decays as the horizon increases with a negligible (median) adjusted $R^2$ of 0.8%, monthly. A regression robust to the impact of outliers, shows a slightly lower predictive power even at the short-term. However, despite this reduction in predictive power, the median adjusted $R^2$ still peaks at 5% at a quarterly horizon. The statistically significance is economically confirmed by computing the maximum attainable unconditional Sharpe ratio that can be reached for each of the models. Across the sample the unconditional buy-and-hold Sharpe ratio is approximately 0.32 annualized. By using, for instance, the predictive regression with the adjusted $R^2$ of 6.82%, the maximal attainable Sharpe ratio would raise to 0.99. The short-term excess returns predictability is robust to the inclusion of other standard predictors, such as the log price-dividend ratio, the log price-earnings ratio, the Term yield spread, the Default premium and the real risk free rate. Table 3 reports the regression results from robust Bayesian regressions. The dependent variable is the one-step ahead excess returns. Robust t-stats are computed by using the standard deviation of the marginal distribution for each of the betas, and reported in parenthesis.

[Insert Table (3) about here]

Columns 1-3 show that, at the very short-term, different measures of the variance premium consistently show one-step ahead predictive power. Columns 4-9 show that lagged implied and realized market variance do not show a significant forecasting ability, once the variance risk premia are included as regressors. Column 10-11 show that, unlike the log price-dividend ratio, the log price-earnings ratio has a statistically significant predictive power, although with the impact is low. The last three specifications add a set of standard predictors to the information set. The predictive role of the variance premium is not crowded out by these additional co-variates, with a statistically significant slope of 0.56 and 0.6 in column 13 and 14, respectively. By including the whole set of predictors the median adjusted $R^2$ raises to a reasonable 7.9%, generating a maximum

\[ (SR_H)^2 = SR_0^2 + \frac{1 + SR_H^2}{1 - R^2} R^2 \]

where $SR_H$ the maximum unconditional Sharpe ratio under the model $H$, $R^2$ the corresponding adjusted R-squared and $SR_0$ the unconditional Sharpe ratio under a buy-and-hold strategy.

The default premium is constructed as the difference between the Baa Moody’s and the long-term government bonds. The term spread is defined as the difference between 10-year and 1-month treasury yields, and the real risk free rate is defined as the difference between the 1-month T-Bill nominal returns and the realized CPI inflation rate not seasonally adjusted.

---

14 I follow Cochrane (1999) linking the maximum unconditional Sharpe ratio attainable using each of the predictive regressions and the corresponding adjusted $R^2$. The mapping is defined as \( (SR_H)^2 = SR_0^2 + \frac{1 + SR_H^2}{1 - R^2} R^2 \) where $SR_H$ the maximum unconditional Sharpe ratio under the model $H$, $R^2$ the corresponding adjusted R-squared and $SR_0$ the unconditional Sharpe ratio under a buy-and-hold strategy.

15 The default premium is constructed as the difference between the Baa Moody’s and the long-term government bonds. The term spread is defined as the difference between 10-year and 1-month treasury yields, and the real risk free rate is defined as the difference between the 1-month T-Bill nominal returns and the realized CPI inflation rate not seasonally adjusted.
attainable Sharpe ratio of 1.07 annualized, against the 0.3 under the buy-and-hold strategy. This Bayesian regression analysis, confirms and extends previous evidence such as those reported in Bali and Hovakimian (2009) and Bollerslev et al. (2009).

2.1 Variance Risk Premium and Macroeconomic Uncertainty

At the outset of the paper, I argue that there may be an economically plausible relation between macroeconomic uncertainty and the variance risk premium. Figure 1 reports the variance risk premium against different empirical proxies for the economy-wide level of uncertainty. As the figure shows, there is a co-movement between (proxies for) uncertainty and the total market variance risk premium. I provide further evidence by regressing the historical variance risk premium on a widely-used set of proxies for economic uncertainty. This set includes, the dispersion of real consumption and GDP one-step ahead forecast, a survey-based index of market uncertainty (see Baker, Bloom, and Davis 2013), the Anxious index, both implied and realized lagged market variances and the predictive variance from a GARCH(1,1) fitted to real per-capita growth rate of consumption. Data are from the survey of professional forecasters held by the Philadelphia Fed and the FREDII database from the St. Louis Fed.\(^{16}\) Table 4 reports the regression results.

As the table shows, macroeconomic uncertainty and the variance premium are positively related. In fact, except for the lagged realized variance, standard proxies for uncertainty such as dispersion of both real consumption and real GDP one-step ahead forecasts, have positive slope coefficients. Interestingly, the equity-related uncertainty measure has a fairly relevant explanatory power, with an adjusted \(R^2\) around 17%. To give a sense of the economic value of this relationship I estimate a main-stream VAR(1) model by including all the variables in table 4. Figure 3 plots the impulse response function of variance risk premium to a shock in the variance of real consumption growth forecasts (top panel), the Anxious index (mid panel), and market uncertainty (bottom panel).

After a shock in the variance of real consumption growth forecasts, the variance risk premium displays a rapid increase of around 2% (on average) within 4 months, and a subsequent rebound from 7 months after the shock. The 1 standard-error interval

\(^{16}\)The Anxious index is a survey-based measure that aims to assess the probability that the GDP will decline in the quarter in which the survey is taken and in each of the following quarters.
highlights that, this hike and rebound is statistically significant at the 5% level. Mid panel shows that, a one-standard deviation shock in the Anxious index at time \( t \), corresponds to an increases in the variance risk premium up to ten months, with a steady decay from 10 to 24 months. Bottom panel shows a similar jump and recovery path after a shock in the index of market-wide uncertainty, although the impact of such a shock is less persistent.

### 3 The Asset Pricing Model

#### 3.1 Preferences

I consider a single-agent discrete-time endowment economy. The investor recursive preferences over aggregate consumption are of the Kreps-Porteus type, allowing for separation between risk aversion and intertemporal elasticity of substitution (Kreps and Porteus 1978, Epstein and Zin 1989, Epstein and Zin 1991 and Weil 1989). This kind of preferences have been also employed by other researchers who studied the relation between macroeconomic fundamentals and the variance risk premium (Drechsler and Yaron 2011, Bollerslev et al. 2009 and Benzoni et al. 2011). Let \( C_t \) denote consumption, the functional form of life-time utility takes the form

\[
V_t = \left\{ (1 - \beta) C_t^{1 - \frac{1}{\psi}} + \beta R_t (V_{t+1})^{1 - \frac{1}{\psi}} \right\}^{\frac{1}{1 - \frac{1}{\psi}}} \tag{2}
\]

in which \( \psi \neq 0 \) is the coefficient of intertemporal elasticity of substitution (IES), \( \gamma \neq 1 \) the relative risk aversion (RRA), and \( \beta \) the subjective discount factor. When the relative risk aversion exceeds the inverse of the IES, the investor prefers early resolution of uncertainty. Let \( P_C^t \) denote the ex-dividend price of a claim on the consumption stream. Epstein and Zin (1989) and Epstein and Zin (1991) show that the SDF can be rewritten as

\[
M_{t,t+1} = \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \beta \frac{P_C^{t+1}}{P_C^t} \right)^{1/(1 - \gamma)} \tag{3}
\]

This is a geometric weighted average between the standard expected CRRA utility component \( \beta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \), and the continuation value of the utility function. The weight assigned to each component of is a function of both RRA and IES. When \( \gamma = 1/\psi \) current consumption growth is sufficient to discount for future payoffs. For \( \gamma \neq 1/\psi \) and \( \psi > 1 \), the higher the level of risk aversion, the higher the weight the agent put on the continuation
utility. Utility is maximized subject to the standard intertemporal budget constraint

\[ W_{t+1} = W_t (1 - k_t) R_{c,t+1} \]

with \( k_t = C_t/W_t \) the fraction of wealth \( W_t \) consumed at time \( t \) and \( R_{c,t+1} \) the gross returns on all invested wealth. The first order conditions take the form

\[
E \left[ M_{t,t+1} R_{c,t+1} | y_t \right] = 1, \quad R_{c,t+1} = \frac{P_{t+1}^C + C_{t+1}}{P_t^C} \quad (4)
\]

\[
E \left[ M_{t,t+1} R_{d,t+1} | y_t \right] = 1, \quad R_{d,t+1} = \frac{P_{t+1}^D + D_{t+1}}{P_t^D} \quad (5)
\]

in which \( R_{d,t+1} \) is the gross return on the market portfolio, and \( P_t^D \) the ex-dividend price of a claim to the asset that pays the dividend stream \( D_t \) for \( t = 1, \ldots, \infty \).

Following Drechsler and Yaron (2011), Bansal and Yaron (2004), Bollerslev et al. (2009), Lettau et al. (2008), Johannes et al. (2011) and Collin-Dufresne et al. (2013) I maintain the assumption that \( \gamma > 1 \) and \( \psi > 1 \). This implies that the representative agent has preferences for early resolution of uncertainty.\(^{17}\) The log return on a one-period risk-free asset, whose value is observed at time \( t \), is defined as \( r_{f,t+1} = -\log \left( E \left[ M_{t,t+1} | y_t \right] \right) \). The equity premium is computed as \( \log \left( R_{d,t+1} - r_{f,t+1} \right) \). Conditional expectations are all taken with respect to the agent’s belief on the growth rate of consumption and dividend. Belief updates acknowledge uncertainty on both state variables and parameters.

### 3.2 Dynamics of Economic Fundamentals

The log real per-capita consumption growth \( \Delta c_{t+1} \) evolves as a time-varying drift plus noise model (West and Harrison 1997 and Harvey 1981)

\[
\Delta c_{t+1} = \mu_{t+1} + \sigma_e \epsilon_{c,t+1} \quad \epsilon_{c,t+1} \sim N(0,1) \quad (6)
\]

with \( \sigma_e \) the idiosyncratic volatility. The expected growth rate of consumption \( \mu_{t+1} \) is a latent AR(1) process with \( \nu \) the level of persistence of the Gaussian and stationary

\(^{17}\)Campbell and Beeler (2012) estimated a value of \( \psi \) well below one. In the model, however, consumption growth is weakly predictable. Therefore, the dynamics of the risk-free rate is almost exclusively due to the belief updating mechanism on the conditional expected growth rate of consumption. As such, a value of \( \psi < 1 \) estimated in a Hall-type regression is not necessarily in contrast with a high level of IES in the actual agent’s preferences.
innovation $\epsilon_{\mu,t+1}$.

$$
\mu_{t+1} = (1 - \nu)E_{\mu} + \nu \mu_t + \sigma_{\mu,\lambda_{t+1}} \epsilon_{\mu,t+1} \quad \epsilon_{\mu,t+1} \sim N(0, 1) \tag{7}
$$

The conditional volatility $\sigma_{\mu,\lambda_{t+1}}$ is time-varying and depends on a two-state Markov regime-switching process where the latent regime $\lambda_t = i$, for $i = H, L$ follows the transition probability matrix

$$
\Pi' = \begin{pmatrix}
p_{LL} & 1 - p_{HH} \\
1 - p_{LL} & p_{HH}
\end{pmatrix}
$$

in which

$$
p(\lambda_{t+1} = H|\lambda_t = H, \theta) = p_{HH} \quad \text{and} \quad p(\lambda_{t+1} = L|\lambda_t = L, \theta) = p_{LL} \tag{9}
$$

and $\sigma_{\mu,\lambda_{t+1}}^2 \in \{\sigma_{\mu,H}^2, \sigma_{\mu,L}^2\}$ with $\sigma_{\mu,H}^2 > \sigma_{\mu,L}^2$. Here, I label $H$ as a state of high macroeconomic uncertainty, and $L$ the opposite. In the transition dynamics $p_{HH}$ $(p_{LL})$ defines the level of persistence of the high(low)-uncertainty state. Notice that, I call $\lambda_t$ the uncertainty state since it drives the conditional dispersion of the agent’s belief on the expected growth rate of consumption, as discussed below. Following Abel (1999), Bansal and Yaron (2004) and Lettau et al. (2008) the aggregate dividend growth $\Delta d_{t+1}$ is defined as a leveraged version of the growth rate of consumption with $\phi > 1$ the rescaling factor.

$$
\Delta d_{t+1} = \mu_d + \phi (\mu_{t+1} - E_{\mu}) + \sigma_d \epsilon_{d,t+1} \tag{10}
$$

Notice, inference on the long-run expected growth rate of both consumption and dividend are not influenced by $\lambda_{t+1}$. In fact, the unconditional mean of consumption growth collapses to $E_{\mu}$, such as $E\left[E_{\mu}|y^T\right]$ coincides with the true unconditional mean as $T \to \infty$. Since $E[\mu_{t+1} - E_{\mu}] = 0$, also the $\mu_d$ coincides with the unconditional mean of the log dividend growth. The exogenous shocks are Gaussian (no jumpi-like shocks) and independent one among the other, i.e. $[\epsilon_c,t+1, \epsilon_{\mu,t+1}, \epsilon_{d,t+1}]' \sim N(0, I_3)$.

### 3.3 Variance Risk Premium

I define the variance risk premium as a direct function of the probability of a high macroeconomic uncertainty state. Let $\mathbb{Q}$ and $\mathbb{P}$ indicate the risk-neutral and the physical probability measure, respectively. From the definition of the variance risk premium and the Radon-Nykodim density ratio $\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{M_{t+1}}{E_{\mathbb{P}}[M_{t+1}]}$, we can map the variance swap rate to
the expectation under the physical measure. Following Carr and Wu (2009) and Miao et al. (2012), the variance risk premium can be computed as

\[ VRP_{t,t+1} = E_t^P \left[ \frac{M_{t,t+1}}{E_t^P [M_{t,t+1}]} RV_{t,t+1} \right] - E_t^P [RV_{t,t+1}] \]

\[ = E_t^P \left[ \frac{M_{t,t+1} RV_{t,t+1}}{E_t^P [M_{t,t+1}]} \right] - E_t^P [RV_{t,t+1}] \]

in which \( E_t^P [M_{t,t+1}] \) the expected value of the SDF under the physical measure. Given the discrete nature of the uncertainty state, by applying the law of iterated expectations, the variance risk premium is calculated as

\[ VRP_{t,t+1} = \tilde{\pi}_{H,t} \times \left( E_{t,\mu}^H [RV_{t,t+1}] - E_{t,\mu}^L [RV_{t,t+1}] \right) + \kappa_t \] (12)

The first component in (12) is convex in \( p(\lambda_{t+1} = H | y_t) \), and bounded at zero when either \( \pi_{t+1}^H = 0 \) or \( \pi_{H,t+1} = 1 \), then peaks when the agent is less confident about which state of the economy is more likely (i.e. \( \tilde{\pi}_{H,t+1} = 0.5 \)). Its functional form is

\[ \tilde{\pi}_{H,t+1} = \left( \frac{E_{t,\mu}^H [M_{t,t+1}]}{\pi_{L,t+1} E_{t,\mu}^L [M_{t,t+1}] + \pi_{H,t+1} E_{t,\mu}^H [M_{t,t+1}]} - 1 \right) \pi_{H,t+1} \] (13)

Clearly the magnitude of this first component is a function of both the subjective probability of a high macroeconomic uncertainty state and the expected value of the pricing kernel conditional on the information available at time \( t \) about both unobservable state variables and parameters. This is a fairly general formulation which allows to investigate the impact of different asset pricing models as far as the discrete nature of the uncertainty state is preserved (see Miao et al. 2012 for an ambiguity-based example).\(^{19}\) Here, specifically, \( E_{t,\mu}^i [M_{t,t+1}] \) represents the expected value of the SDF under recursive preferences at time \( t \), given the \( i \)th uncertainty regime and the drift \( \mu_{t+1} \). The second component is always positive since \( Var_{t,H} [\mu_{t+1}] > Var_{t,L} [\mu_{t+1}] \) generates \( Var_{t,\mu}^H [RV_{t,t+1}] > Var_{t,\mu}^L [RV_{t,t+1}] \) under recursive preferences for early resolution of uncertainty. The third component \( \kappa_t \) is convex in \( \tilde{\pi}_{H,t+1} \) and is positive since \( Cov_{t,\mu}^L [RV_{t,t+1}, M_{t,t+1}] > 0 \) although numerically negligible (see the appendix). I leave more complete analytical expressions to the

\(^{18}\)Notice here the shocks on fundamentals are all Gaussian and independent one among the other. As pointed out in Drechsler and Yaron (2011) this makes the variance spread exclusively determined by the difference of the expected values under different measures of the same quantity, which is the total variance of market returns.

\(^{19}\)Notice that, because of real-time structural learning, the conditioning information set radically changes, generating different pricing implications with respect to Miao et al. (2012). From this perspective their work, which in turns echoes the decomposition in Carr and Wu (2009), may be seen as a general setting.
3.4 Real-Time Learning and Macroeconomic Uncertainty

Conventional wisdom and most asset pricing research, posits that individuals fully observe both the state variables and the parameters that govern the dynamics of the economy. Those works that depart from such a strict assumption focus on learning either the structural parameters or state variable, alternatively.\(^{20}\) This paper differs from existing research along one key dimension. I focus on the empirical implications of simultaneously learning about parameters and state variables.

In the model, the agent only observes \(y_t = (\Delta c_t, \Delta d_t)\), and jointly learns in real-time both the nature of the latent states \(z_t = (\mu_t, \lambda_t)\), and the vector of structural parameters \(\theta = (E\mu, \nu, \mu_\phi, \phi, \sigma^2_\mu, \sigma^2_\phi, p_{LL}, p_{HH}, \sigma_L, \sigma_H)\), fully acknowledging their uncertain nature. Learning about multiple unknowns is not only more difficult, as additional unobservables confound inference, but also allows to relax some of the crucial assumptions usually needed to generate sensible implications in standard asset pricing settings, such as high persistence of exogenous shocks in the dynamics of states variables (see Williams 2003 and Carceles-Poveda and Giannitsarou 2008 for a more detailed discussion). Indeed, the existence of such persistent components is still under debate (see Hansen 2007, Hansen et al. 2008, Hansen and Sargent 2010, and Sargent 2007 among the others).\(^{21}\) More generally, although may be detectable ex-post, a high level of persistence in, say, the conditional volatility of consumption growth, is not necessarily perceivable ex-ante. Given the forward looking nature of the Euler equations, unless the investor observes at each time \(t\) the low-frequency component in consumption growth, it is not clear how the shocks could be reflected in equilibrium prices.

Without having to rely on such a strong assumption, parameter learning endogenously generates low frequency shocks in the conditional distribution of the expected growth rate of future consumption. Simply put, agent’s belief updates are non-stationary. For example, if the agent realizes that the unconditional growth rate of consumption \(E\mu\) is higher than previously thought, the belief revision mechanism generates a permanent

---

\(^{20}\)Learning about a hidden states falls under the heading of signal extraction. The structural parameters of the economy are fixed and the underlying state of the economy, for instance the business cycle, is extracted from observable consumption/dividends. In contrast, learning about structural parameters boils down to parameter uncertainty.

\(^{21}\)In particular Hansen et al. (2008) point out that: “Many of the statistical challenges that plague econometricians presumably also plague market participants. Naïve application of rational expectations equilibrium concepts may endow investors with too much knowledge about future growth prospects. Learning and model uncertainty are likely to be particular germane to understanding long-run risk.”
shock which is directly reflected in the conditional distribution of future consumption growth. This “random-walk” property of parameter learning generates a particular strong form of long-run risk which is heavily priced under Kreps-Porteus preferences (see also Collin-Dufresne et al. 2013 for a related discussion).

Real-time learning about state variables and parameters occurs in two steps; at time $t$ the agent holds joint beliefs over the latent states and parameters $p(\theta, z_t | y_t) = p(z_t | \theta, y_t) p(\theta | y_t)$ with $y_t = (y_1, ..., y_t)$ the history of consumption and dividend growth rates. Thus, the investor first computes the predictive distribution given current information

$$p(z_{t+1}, \theta | y_t) = \int p(z_{t+1} | z_t, \theta) p(z_t, \theta | y_t) dz_t$$  \hspace{1cm} (14)

Second, updates her beliefs given the predictive likelihood $p(y_{t+1} | z_{t+1}, \theta)$

$$p(z_{t+1}, \theta | y_{t+1}) \propto p(y_{t+1} | z_{t+1}, \theta) p(z_{t+1}, \theta | y_t)$$  \hspace{1cm} (15)

This two-steps procedure show the recursive nature of Bayesian updating. In fact, $p(z_{t+1}, \theta | y_{t+1})$ is functionally dependent on $p(z_t, \theta | y_t)$. The sequential nature of structural learning makes the posterior at time $t + 1$ the prior for date $t + 2$ and so on. As such, the main issue is characterizing $p(z_t, \theta | y_t)$ for each time $t$. Unfortunately, there is not a natural way to introduce parameter learning in a well-posed manner getting easily interpretable closed form approximations. Even though $\lambda_t$ is discretely valued, there is not analytical form for $p(z_t, \theta | y_t)$, as it is high-dimensional and the dependence on the data is highly complicated and nonlinear. This learning mechanism leads to a rapidly increasing curse of dimensionality as the sample information cumulates. I solve the computational burden by using a particle filtering and learning scheme (Carvalho, Johannes, Lopes, and Polson 2010b, and Carvalho, Johannes, Lopes, and Polson 2010a).

Macroeconomic uncertainty is defined as the dispersion of the agent’s belief about the expected growth rate of future consumption given current information

$$E \left[ (\mu_{t+1} - E[\mu_{t+1} | y_t])^2 | y_t \right]$$

with $E[\mu_{t+1} | y_t]$ computed from the predictive distribution $p(\mu_{t+1} | y_t)$. This predictive is
obtained by integrating out uncertainty on both parameters and the state $\lambda_{t+1}$, which makes it heavy-tailed, then leading the agent to see the economic outlook as riskier.\footnote{In a recent paper Ludvigson, Jurado, and Ng (2013) used a similar definition of macroeconomic uncertainty. In their framework, however, neither states nor parameters uncertainty is considered whatsoever. This would make the predictive distribution Gaussian.}

Under the agent’s filtration, macroeconomic uncertainty is defined as

$$Var[\mu_{t+1}|y^t] = \nu^2 C_t + \sigma^2_{\mu}$$

where $C_t = Var[\mu_t|y^t]$ the posterior variance of $\mu_t$, $\nu$ the persistence parameter and $\sigma^2_{\mu}$ the conditional variance of the expected growth rate marginalized over the uncertainty state (see West and Harrison 1997 and the appendix). Notice that, macroeconomic uncertainty is increasing in both the persistence parameter $\nu$, and the probability of a high uncertainty regime. Higher persistence makes the impact of current belief dispersion bigger. Similarly, the more likely is $\lambda_{t+1} = H$ the higher the economic uncertainty. This may be easily understood from the dispersion of the agent’s belief on the expected growth rate of consumption conditional on the state $\lambda_{t+1}$,

$$Var[\mu_{t+1}|\lambda_{t+1} = i, y^t] = \nu^2 C_t + \sigma^2_{i,\mu}$$

from here, macroeconomic uncertainty is computed as the marginal dispersion over the regimes (predictive) probabilities

$$Var[\mu_{t+1}|y^t] = \sum_{i=1}^{k} p(\lambda_{t+1} = i|y^t) \times Var[\mu_{t+1}|\lambda_{t+1} = i, y^t]$$

(16)

Now, since by construction $\sigma^2_{H,\mu} > \sigma^2_{L,\mu}$, the higher the probability of a high uncertainty state, the higher the level of macroeconomic uncertainty, such as jumps in $p(\lambda_{t+1} = H|y^t)$ generate “uncertainty shocks” which are, within the model, directly related to the variance risk premium.

4 Empirical Analysis

I now briefly discuss the parameters estimates of the dynamics depicted in Section 3.2, then I show the equilibrium asset pricing implications. Despite a low persistence of the conditional mean and volatility of consumption growth, the model is able to quantitatively match a broad set of unconditional moments of both the variance risk premium and equity returns, as well as replicate excess returns predictability.
4.1 Parameters Estimates

Unlike the reference literature the model parameters are estimated in real-time fully acknowledging their uncertain nature. After setting the prior hyper-parameters, Bayesian learning evolves naturally across the sample, which is 1990:01 - 2013:01, monthly. I leave more details on the prior specification to the appendix. The first 4 years of monthly estimates are cut as a burn-in sample to get rid of prior dependence in the equilibrium results. The marginal posterior mean of each parameter are is computed at each time \( t \) as

\[
E[\theta | y'] = \int \theta p(\theta, z_t | y') d\theta dz_t
\]

with \( y' = (y_1, ..., y_t) \) the available information and \( y_r = (\Delta c_r, \Delta d_r) \). These are marginal beliefs, as uncertainty about the state \( z_t \) is integrated out. Similarly the posterior mean of the states \( z_t \) is computed as

\[
E[z_t | y'] = \int z_t p(\theta, z_t | y') d\theta dz_t
\]

Panel A in Figure 4 shows the real-time estimates of the persistence parameter \( E[\nu | y'] \), together with the corresponding 95% confidence region. The agent’s belief on the persistence of the expected growth of consumption steadily decreases up to late 2000s, where the posterior mean reaches a lower bound of 0.1.

The zero value falls within the 95% confidence interval throughout the 2000s. This means that, from 2001 to 2008, the conditional mean of consumption growth is seen by the agent as a constant and equal to \( E[\mu] \). As such, the growth rate of consumption is perceived as an i.i.d process, conditional on the Markov regime-switching state \( \lambda_{t+1} \). After the great financial crisis the posterior mean of \( \nu \) slightly increases with an end-of-sample estimate equal to \( E[\nu | y^T] = 0.2 \). Such a low level of persistence, implies a fairly weak predictability in consumption growth, consistent with Campbell and Beeler (2012). Panel B shows that the posterior mean of the long-run growth rate \( E[E[\mu] | y'] \) is highly persistent and low volatile. However, this high persistence does not imply predictability in consumption growth since \( E[\mu] \) is constant in the data generating process, and time variation is only due to the learning mechanism. These two pictures together show that, from an ex-ante perspective, the dynamics of the expected growth rate of consumption is not perceived as persistent as usually assumed in the existing literature. This is also
true with respect to the dynamics of the conditional volatility of economic growth. Top panel in Figure 5 show the perceived probability of a high uncertainty state.

Spikes in uncertainty are relatively infrequent and located across the end of the 1990s (LCTM/Russian Crisis), early 2000s (Worldcom, Enron and 9/11 attacks) and the recent great financial crisis of 2008/2009. In the model, these spikes increase the level of macroeconomic uncertainty, which is important for capturing not only the variance risk premium but also the higher moments in the unconditional distribution of market returns.

As shown in the figure, under the agent’s filtration shocks in the conditional volatility of consumption growth are infrequent (but not rare), and relatively large (compared to standard Gaussian shocks, because of their discrete nature). The persistence of these volatility shocks is relatively low. However, each spike in the conditional volatility has an effect on the belief updating mechanism. As a matter of fact, the higher the conditional volatility of the expected growth rate of consumption, the higher is the weight put by the agent’s on current growth rate of consumption. As the agent puts more weight on information from fundamentals, the posterior means of parameters change as well. Therefore, even though transitory, shocks in volatility have a persistent effect on the equilibrium conditions through parameter learning.

The model is solved on a real-time basis and is fed with sequential estimates of state variables and parameters at each time $t$. For the sake of completeness I show in the appendix also the mean, median and 95% percentiles of the end-of-sample parameter estimates of the dynamics of both consumption and dividend growth. These values represent those estimates obtained by the agent once the entire sample of consumption and dividend is observed. I use the end-of-sample estimates to calibrate the fixed-parameter model specification, in order to disentangle the role of structural learning from states filtering. This may be interpreted as a rational expectations benchmark with no parameter uncertainty.

Tauchen (2012) and Bollerslev et al. (2009) show that the main channel through which persistence effectively has an impact on asset prices is towards shocks on the conditional volatility of consumption growth.

They may be thought as standard maximum likelihood estimates, provided learning is unbiased, which is the case here.
4.2 Asset Pricing Results

A consistent characterization of real-time learning comes at a cost. The representative agent maximizes life-time utility at each time $t$ based on joint posterior beliefs on states and structural parameters. These are updated from current consumption and dividend growth. Therefore, both the optimal wealth-consumption and the equilibrium price-dividend ratio become belief- and time-dependent, such as close form solutions become unfeasible. In the model, there are ten parameters and two unobservable states governing the dynamics of economic fundamentals. The agent’s belief for each of those is governed by two hyper-parameters accordingly, further introducing nuisance parameters, for a total of twenty-four unknown in the Euler equations. This, unfortunately makes the state-space prohibitively large. In order to keep the numerical solution feasible, I recursively solve the equilibrium model by using an anticipated utility approach (see also Kreps 1998, Piazzesi and Schneider 2010, and Cogley and Sargent 2009). The key idea is that, ex-ante, beliefs on states and parameters are seen as if they will remain constant indefinitely into the future, albeit ex-post they change over time due to learning. This method makes standard numerical methods applicable, maintaining the time series implications on the model dynamics inherited from parameter learning. The appendix explains how I solved numerically the equilibrium model by using posterior mean beliefs and a grid of values for the beliefs on the state variables.

Figure 6 shows both the historical and the model-implied dynamics of the variance risk premium. As far as the unconditional moments are concerned, Tables 5-9 provide the empirical results and the corresponding statistics from the estimated model, for both the variance and the equity risk premium. The sample period is 1990:01 - 2013:01, and the first 4 years of monthly results are cut as a burn-in sample. In comparing the model with the data, I show the mean, median and 95% percentiles generated by simulating 20,000 returns at each time $t$. Standard errors and confidence intervals for the historical estimates are generated via non-parametric bootstrap. I provide results for different values of relative risk aversion ($\gamma = 2, 5$), with the intertemporal elasticity of substitution $\psi = 3.5$. Further results with $\psi = 1.5$ are provided in a separate on-line appendix.

4.2.1 Variance Risk Premium. Figure 6 reports the historical vs model-implied dynamics of the variance risk premium. The red line shows the historical market premium for the variance risk which is measured as in section 2. The blue line represents the expected value at each time $t$ of the variance risk premium under the model.

[Insert Figure (6) about here]
Despite a bit of misalignment at the end of the testing sample, it is fair to say that the model is able to match reasonably well the conditional dynamics of the variance risk premium. This is especially evident across specific periods such as late 1990s, early 2000 and the across the recent great financial crisis. These periods coincide with the shocks in the level of economic uncertainty depicted in figure 5. Notice that, remarkably, the model-implied variance risk premium endogenously raises uniquely on the basis of macroeconomic fundamentals, and no information from prices, or financial markets in general, are used.

Table 5 reports the statistics for the variance risk premium. I first focus on the results from the model with real-time structural learning in comparison with a rational expectations benchmark, then I discuss the results from a model with real-time structural learning and CRRA preferences. The aim is to disentangle the specific role that structural learning and recursive preferences may play in explaining the unconditional moments of the variance risk premium, and its short-term excess returns predictive ability. Panel A show the results statistics.

[Insert Table (5) about here]

The model is able to generate a sizable average variance risk premium consistent with the data. Hence, model-implied index options contain a valuable insurance premium. Panel A also shows that the model replicate the large unconditional volatility of the variance premium. This reflects the impact of structural learning on both the perceived probability of a high uncertainty state and the gap between the market return variance between high and low uncertainty regimes, as discussed in Section 4.3. The table further shows that the model with real-time structural learning also matches the high positive skewness and excess kurtosis of the market variance risk premium, with the historical values fall within the 95% model-implied confidence intervals. In the model, spikes in macroeconomic uncertainty generate a higher mass of probability on the right tails of the unconditional distribution of the variance risk premium, such as the unconditional distribution of the mode-implied index options premium depart from normality, consistent with the data.

The results computed from the rational expectations benchmark are reported in the middle panel of Table 5. By excluding parameter learning, both the model-implied unconditional mean and volatility of the variance risk premium are sensibly lower than in the data. Therefore, uncertainty on economic fundamentals actively increases the premium the agent is willing to pay to hedge for market variance positive shocks, as the agent does not directly observes the transition dynamics between uncertainty states. Additional unknowns makes inference on the uncertainty state more difficult, increasing the
perceived probability of a positive shock in the conditional market variance. This result is consistent with Miao et al. (2012). They show that (partial) incomplete information may not be enough to explain the variance risk premium under recursive preferences.

Bottom panel in Table 5 shows further evidence that parameter uncertainty may be interpreted as an extra-source of risk which inflates the equilibrium variance risk premium. The table reports the results computed from a model with CRRA preferences in which the agent no longer cares about the continuation value of the life-time utility as with recursive preferences. This feature of CRRA preferences makes the agent less willing to put more weight on the pricing kernel during regimes of high macroeconomic uncertainty. However, parameter uncertainty makes the economy riskier as the expected growth rate of consumption is no longer conditionally Gaussian distributed. This, together with the impact on the subjective probability of a high state of economic uncertainty discussed above, counterbalance the lower volatility on the pricing kernel inherited from CRRA preferences. As the table shows, around half of the unconditional mean and volatility of the variance risk premium can be replicated by means of real-time structural learning. Also, the historical positive skewness still falls within the 95% confidence interval implied by the model. Nevertheless, the role of recursive preferences is still crucial. In fact, neither the unconditional mean and volatility of the index options premium effectively coincide with those implied by the historical information.

Table 5 shows, as a whole, two results concerning the coefficient of relative risk aversion. First, the model matches unconditional properties of the variance premium with a relatively low level of risk aversion ($\gamma = 5$). This value is sensibly lower than some existing research, such as Drechsler and Yaron (2011) and in line with Drechsler (2013) and Miao et al. (2012). Second, the unconditional moments of the variance premium are relatively stable slightly decreasing the investor’s risk aversion. This lower effect of the relative risk aversion is due to the fact that, while decreasing $\gamma$ from 5 to 2 slightly decreases the first component in (13), the second component is almost unaffected as the gap between the conditional market variance between the high and low uncertainty regimes is relatively stable across different levels of risk aversion.

It is worth to mention that, a large variance risk premium is by no means a direct consequence of the fact that the model also matches the first two unconditional moments of equity returns. Indeed, as discussed in a separate appendix, a standard calibrated long-run risk model, even though matches the mean and volatility of equity returns, generates a counter-factual constant and numerically negligible variance risk premium. A separate argument concerns the role of higher order moments. Since skewness and excess kurtosis have an important effect on index option prices, it is important that the
model also matches their historical values, as discussed in the next section.

Section 2 highlights the short-term predictive ability of the variance premium for historical market excess returns. This predictive power is robust to the inclusion of standard long-run predictors, such as the log price-dividend ratio. In the model, economic uncertainty shocks coincide with a spikes in the variance risk premium and a drop in the price-dividend ratio then a decreasing realized returns. This in-sample negative correlation becomes positive out-of-sample because of the mean reverting nature of the agent’s belief on the expected growth rate of consumption, which translates in mean reverting equity returns. Thus, ex post the econometrician observes that a relatively high level of variance risk premium is followed by increasing market excess returns. The mechanism is opposite for the log price-dividend ratio, which is positively in-sample correlated with equity returns (see Timmermann 1993 and Lewellen and Shanken 2002 for a related discussion). The predictive horizon depends on the persistence of the predictors. The model-implied variance premium is much less persistent than the log price-dividend ratio. As such, out-of-sample predictability tends to die out quicker. Table 6 provides the statistics results of

\[
\frac{1}{k} \sum_{i=1}^{k} (r_{m,t+k} - r_{f,t+k}) = \alpha_k + \beta_{VRP}^{k} VRP_{t,t+1} + \beta_{lpd} lpd_t + u_{t,t+k} \quad k = 1, \ldots, 24
\]

This predictive regression is run with overlapping monthly returns. The dependent variable is the historical equity premium \(r_{m,t+k} - r_{f,t+k}\), averaged over the forecasting horizon. The variance premium \(VRP_{t,t+1}\) and the log price-dividend ratio \(lpd_t = \log(P_D_t)\), implied by the model, are the independent variables. Regressions are run for each of the 20,000 simulated outcomes from the model, generating median slope coefficient estimates, as well as median robust t-stats and adjusted \(R^2\). The t-stats are corrected for heteroskedasticity and autocorrelation in the residuals.25

Top panel of Table 6 shows that the model also quantitatively captures the short-term predictive ability of the variance risk premium. As the table reports, the average slopes, and corresponding t-stats, of the variance premium are decreasing as the forecasting horizon increases. On the other hand, the absolute value of the slope coefficients on the log price-dividend ratio, and related t-stats, are increasing with the predictive horizon. These

\[\text{[Insert Table (6) about here]}\]

\[\text{[25Here I do not run Bayesian regressions. As a matter of fact, since the outcome is 20,000 simulated betas, t-stats and adjusted } R^2, \text{ a further Bayesian analysis would be redundant, at least for the explanatory purpose of this section.}\]
results are consistent with the data and previous evidence in Section 3. Furthermore, the median adjusted $R^2$ is sensibly increasing with the forecasting horizon, with a value of 2% at the one-month horizons that increases to 10% at the one-year horizon.

Panel B shows the results from the rational expectations benchmark. Let recall this benchmarking model is calibrated by using the end-of-sample parameter estimates from the model with Bayesian learning on state variables and structural parameters. The variance premium slope is no longer statistically relevant at the short-term and the slope of the log price-dividend is significant only at the longest horizon ($k = 12$). However, the average adjusted $R^2$ is both numerically and economically negligible. Without the endogenous low-frequency shocks inherited from parameter learning, the model does not generate excess returns predictability. The high stationarity of the exogenous innovations does not generate enough in-sample variation of both the variance risk premium and the log price-dividend ratio to explain the future excess returns. In other words, the signal-to-noise ratio in these regressions, meaning the ratio between the variance of the variance premium and the log price-dividend ratio with respect to the market excess returns, is too low to result in significant predictability in small-samples. In contrast, the higher variation of both the model-implied variance premium and log price-dividend ratio, sensibly increases the signal-to-noise ratio, then raising the regression $R^2$.

Without persistence in the state variables, the model does not show any kind of predictive ability. To better investigate this finding, and specifically if parameter uncertainty may effectively substitute exogenous persistence, I also compute the same set of predictive regressions by using the model-implied variance risk premium computed from the rational expectations benchmark, in which however I now exogenously impose an high persistence dynamics of the conditional expected growth rate of consumption. Panel C reports the results. As the table shows, by increasing the impact of stationary shocks on the dynamics of fundamentals, the model generates a slightly stronger predictive power. The slope coefficient of the variance premium peaks at the quarterly horizon, and the beta on the log price-dividend ratio increases in magnitude as the forecasting horizon increases, although the median adjusted $R^2$ is lower than with parameter uncertainty. To summarize, parameter learning generates excess returns predictability, consistent with Lewellen and Shanken (2002). In addition, an ad-hoc rational expectations benchmark shows that predictability is positively linked with the persistence of the latent states dynamics, consistent with Bollerslev et al. (2009).\footnote{They show that both the magnitude of the slope and the $R^2$ are positively related to the level of persistence of the state variables.}

26
4.2.2 Cash Flows and Equity Returns. Table 7 shows that the model matches reasonably well most of the key unconditional moments of both consumption and dividend growth. Here, I report the means, medians and 95% confidence intervals for 20,000 simulations. Cash-flows are simulated at each time $t$ from their marginal distributions $p(\Delta c_t | y^t)$ and $p(\Delta d_t | y^t)$. Simulations results are computed monthly, then aggregated annually.

[Insert Table (7) about here]

The historical moments fall within the model-implied 95% confidence intervals, and the median estimates are relatively close to the data. The model matches also higher moments in the cash-flows unconditional distributions, such as negative skewness and excess kurtosis. This is a by-product of the unconditional distribution of the agent’s belief on the expected growth rate of consumption. As shown in Figure 7, this belief is negatively skewed and fat-tailed.

[Insert Figure (7) about here]

As a matter of fact, the regime-switching nature of the latent state $\lambda_t$ and the presence of parameter uncertainty, make the marginal belief on the economic growth rate heavy-tailed. As discussed above, time-varying uncertainty makes the agent over-weights (under-weights) observations during periods of high (low) economic uncertainty (see Van Nieuwerburgh and Veldkamp 2006). As such, belief revision in bad times are deeper than under good times (see Veronesi 1999). This makes structural learning asymmetric, increasing the agent’s perception of a tail risk associated with the growth rate of the economy. Macroeconomic uncertainty shocks command then a high risk premium as in weak economic conditions stock prices drop, they might drop heavily, increasing the return variance. These price, tail and variance risk together are jointly embedded in the shape of the agent’s unconditional belief in Figure (7). Finally, the fact that the agent’s belief is leptokurtic means that uncertainty shocks increase to increasing concerns about extreme events, skewed towards worries about bad (or disastrous) events.

Table 8 shows the model implications for aggregate market returns. The table reports both the model-implied statistics as well as the annual historical information on equity returns, the log price-dividend ratio and the real risk-free rate, with corresponding bootstrap standard errors. The historical equity premium is 4.92% and the returns volatility is around 17.05%. The data shows a significant negative skewness, (i.e. -0.9), and high excess kurtosis (i.e. 4.07). The log price-dividend ratio is highly persistent, with a first order autocorrelation coefficient equal to 0.94. The model matches the un-

\[^{27}\text{Note this marginal are not Gaussian distributed since uncertainty on both state variables and parameters have been integrated out.}\]
conditional moments of market returns for a reasonably low level of risk aversion ($\gamma = 5$) and leveraging factor ($\phi = 2.62$). At the same time, the model also matches the negative skewness and excess kurtosis, which characterize the unconditional distribution of market excess returns. This is particularly important as a robustness check on the variance risk premium unconditional moments. Indeed, there is a significant relation between higher moments of market returns and the premium embedded in index options, as shown in Bakshi and Madan (2006).\textsuperscript{28} This feature of the model is not drawn by leading asset pricing models such as the long-run risk (see Bansal and Yaron 2004, Bansal, Kiku, and Yaron 2007 and results provided on a separate on-line appendix), and Habit-Formation (see Campbell and Cochrane 1999). The mechanism that generates higher moments in the unconditional distribution of equity returns is the same already discussed for the dynamics of cash-flows. Time-varying macroeconomic and parameter uncertainties generate both negative skewness and excess kurtosis.\textsuperscript{29}

Table 8 further shows that, both mean and volatility of the real risk-free rate fall within the 95% model-implied confidence intervals. In addition, the log price-dividend ratio is highly persistent, similar to the data, though its unconditional mean (volatility) is a bit low (high), with $\gamma = 5$, and both slightly higher than the data with $\gamma = 2$.

Panel B in Table 8 shows asset pricing unconditional moments from the rational expectations benchmark. Without parameter uncertainty, and given the low persistence implied by the end-of-sample estimates calibration, the model is not able to replicate most of key unconditional properties of the aggregate market returns, the real risk-free rate and the log price-dividend ratio. Especially the log valuation ratio turns out to be low volatile and low persistent, contradicting the empirical evidence. The low persistence is a consequence of the model’s calibration and the absence of parameter uncertainty. As a matter of fact, while with parameter learning the average autocorrelation of the log price-dividend ratio is 0.96, the rational expectations benchmark shows a significantly lower 0.23. In order to check for this endogenous long-run risk property of parameter learning, I also compute the whole set of unconditional moments calibrating the rational expectations benchmark along the line of Bansal and Yaron (2004), imposing $\nu = 0.98$. This would replicate

\textsuperscript{28}In particular, Bakshi and Madan (2006) show that the premium on market index options is positively linked to the presence of excess kurtosis and negative skewness in the unconditional distribution of stock market returns.

\textsuperscript{29}Remember the agent does not observe the structure of the exogenous innovations. As such, under the agent’s filtration, the error terms in the data generating process have correlation equal to one. In addition, in the data a positive shock in macroeconomic uncertainty is more likely to be negatively related to the expected growth rate of consumption.
a standard long-run risk framework with incomplete information about the underlying states \( z_t = (\mu_t, \lambda_t) \). Panel A in Table 9 reports the results.

[Insert Table (9) about here]

As shown in the table, by increasing the level of persistence in the conditional expected growth rate of consumption, the model performances in terms of matching the unconditional moments, sensibly improve. This is a fairly interesting result, which points out how parameter uncertainty may effectively generate a long-run type of risk, which is not necessarily perceived ex-ante, although may be detectable ex-post by looking at the time series of consumption growth.\(^{30}\)

As a robustness check, I finally investigate the specific role of parameter learning by computing the unconditional moments of returns within a CRRA setting and real-time parameter learning. Panel B of table 9 shows the results. The model with \( \gamma = 5 \) explains on average around 45\% of the aggregate equity premium, as well as a fairly relevant fraction of market returns volatility. However, the model ability to explain key higher order moments of equity returns reduces. Although the historical excess kurtosis falls within the model-implied 95\% confidence interval, the level of skewness is numerically negligible, albeit slightly negative on average. As a whole, however, Table 8 and 9 show that parameter uncertainty may generate a similar impact as persistent state variables, with the sensible advantage of generating the same sort of asset pricing implications – at least in terms of sign of higher order moments – even within a standard CRRA setting.

### 4.3 Robustness Checks and the Pricing Mechanism

In the model, the variance risk premium depends on both the hidden state of macroeconomic uncertainty and incomplete information on structural parameters. Upward shocks in macroeconomic uncertainty increase the variance risk premium, while the presence of parameter learning inflates the impact of these shocks because of both its effect on the agent’s information set and on the equilibrium pricing kernel. The positive relation between macroeconomic uncertainty and the variance risk premium is driven by the common latent regime \( \lambda_t \). Indeed, the definition in (16), implies that the higher the probability of the high-uncertainty regime, the higher the dispersion of the agent’s belief on the expected growth rate of consumption. As shown in figure 8, this positive link is

\(^{30}\)For the sake of clarity I must say that Collin-Dufresne et al. (2013) already mentioned the potential source of long-run risk embedded in parameter uncertainty. Here I confirm and extend their evidence by showing that parameter learning effectively compensate the absence of persistence in the dynamics of state variables.
confirmed by the data,

Top panel shows the model-implied relation between the probability of a high uncertainty state and the level of macroeconomic uncertainty, which is positive by construction. Bottom panel shows the relation between the same model-implied probability against the predictive variance from a GARCH(1,1) fitted on the real growth rate of consumption. The more likely is the uncertainty state, the higher is the predictive variance of consumption growth, which is commonly seen as a proxy for economic uncertainty. The model implies that a shock in the uncertainty state corresponds to an increase in the variance risk premium. Therefore, from Figure 8 it would be sensible to expect a positive relation between macroeconomic uncertainty and the index options premium. Figure 9 show that this link is confirmed in the data

Bottom panel reports the positive relation between the predictive variance of consumption growth from a GARCH(1,1) and the historical market variance risk premium. This confirms the model implications, which is reported in the top panel. These evidence confirm the empirical results I present at the onset of paper.

As discussed above, the model implies that there is a compounding effect of time-varying macroeconomic uncertainty and parameter learning. The specific role of real-time structural learning may be tricky to be investigated. As a matter of fact, although the model implies that the market variance risk premium is driven by subjective belief on the uncertainty state, we can not a priori rule out the potential effect of macroeconomic risk and drops in the expected growth rate of consumption, as they affect the higher order moments of equity returns. If the model is not mispecified it would be sensible to expected that belief updates on the uncertainty state and the conditional moments of consumption growth affect changes in the variance risk premium. I test this assumption by regressing changes on the variance premium $VRP_{t,t+1} - VRP_{t-1,t}$ on belief updates about the uncertainty state $p(\lambda_{t+1} = H|y^{t+1}) - p(\lambda_{t+1} = H|y^t)$.\footnote{I'm using the revision of belief instead of differences in predictive since the predictive at time $t+2$ is functionally related to the posterior at time $t+1$ via the transition matrix $\Pi$, (see Hamilton 1994).} I also add as further controls belief updates on the conditional expected growth rate of consumption $E[\Delta c_{t+1}|y^{t+1}] - E[\Delta c_{t+1}|y^t]$ and macroeconomic risk $Var[\Delta c_{t+1}|y^{t+1}] - Var[\Delta c_{t+1}|y^t]$.\footnote{I use the end of period time measurement convention for the real per capita consumption. As pointed out in Campbell and Viceira (1999) by using beginning or end of time period may not generate different results, in qualitative terms, given the time-averaging nature of the consumption measurement.}

Both the conditional expectations $E[.|y^t]$ and the conditional variances $Var[.|y^t]$ are
computed integrating out parameter uncertainty. These two belief revision are included to check for the impact of a lower than expected consumption growth, and an higher than expected macroeconomic risk, respectively. Finally, I also include current and past consumption growth. By controlling for these variables, I ensure that the slope are driven by the belief revision process, and not by past information available. Both dependent and independent variables are rescaled by their standard deviations such that a one percent increases in the belief revision on the uncertainty state is associated with, for instance, a 0.18% positive changes in the variance premium.

Table 10 does not reject the null of a positive relation between macroeconomic uncertainty shocks and the variance risk premium. Column 1, for instance, shows that a positive belief updates on the uncertainty regime, corresponds to a 0.19% increase in the market variance risk premium. Belief revision on the conditional expected growth rate has a negative effect, with an average -0.1% across columns. The mechanism is the following. Since the uncertainty state is counter-cyclical, while consumption growth is the opposite, a negative revision on the expected growth rate of future consumption may increases the fear of unfavorable shocks to the investment opportunity set due to higher volatility.

In contrast, belief updates on the conditional consumption risk have a positive effect on index options, carrying an average 0.2% positive time spread. The underlying mechanism is the same as discussed above. An higher than expected economic risk is positively related to a positive uncertainty shock. This, in turn, generates an increase in the variance premium. These results hold controlling for both past and current real consumption growth. This is fairly strong results which, as far as I am aware of, has not been previously mentioned in the variance risk premium literature. The adjusted $R^2$ spans from a relatively low 5.3% by including only the belief revision about the uncertainty state, to a fairly high 14.8% reached by including all belief revisions.

In order to investigate the specific role of parameter uncertainty on the variance risk premium, Figure 10 shows the first component in (12), computed from the model with

---

33 Here, I do not include revision in belief within the fixed-parameter case as further controls since the goal here is to check for the role of belief revision as a whole. Johannes et al. (2011) show that the role of parameter learning is robust to the inclusion of fixed-parameters revision.

34 The fact that consumption growth and uncertainty are negatively correlated has been already pointed out in the literature. Higher uncertainty can induce households a buffer stock of savings, then reducing consumption expenditures.
real-time learning and from the rational expectations benchmark.

Parameter uncertainty does seem to sensibly contribute to generate a large variance risk premium. In fact, with parameter learning the preferences-adjusted future probability of a high-uncertainty state is sensibly higher than under rational expectations, since the agent is not only uncertain about the underlying state of the economy, but also about the transition mechanism driving the dynamics of this state. To illustrate, suppose that the investor realizes that the transition probability from a state of low to high macroeconomic uncertainty is higher than previously thought. This generates a wider dispersion in the future probability of a high economy-wide uncertainty regime which does not come directly from the regime itself, increasing the agent’s subjective belief of a high uncertainty state. Figure 11 shows the gap between the probability of a high uncertainty state within real-time parameter learning and the rational expectations benchmark.

The difference is one average positive and hikes across the great financial crisis. Intuitively this affects the market variance risk premium in equilibrium, leveraging up the difference between the conditional market variance in the high and low uncertainty regimes. Another results from the model is that unconditional moments of the market variance risk premium are relatively stable across different values of relative risk aversion. This may sound counter-intuitive, since, within recursive preferences, the higher the level of risk aversion, the higher the weight the agent puts on the pricing kernel in a regime of high macroeconomic uncertainty, increasing the model-implied equilibrium variance premium. However, as shown in Figure 12, the impact of the coefficient of risk aversion on the first component in (12) is relatively low.

Therefore, the impact of relative risk aversion on the market variance risk premium is apparently flawed, consistent with the results statistics.
5 Conclusion

This paper argues that positive shocks in macroeconomic uncertainty – which is defined as the dispersion of a single agent belief on the expected growth rate of consumption – may help explain the large and volatile premium embedded in equity index options, the so-called variance risk premium. I first document this fact empirically by using several proxies for economic uncertainty, showing that upward movements in macroeconomic uncertainty can be connected to an increasing variance premium. I then study this positive link within a general-equilibrium endowment economy, in which a single agent has Kreps-Porteus recursive preferences and learns in real-time the structural parameters governing the dynamics of economic fundamentals.

I show that infrequent, large, and transitory uncertainty shocks generate a sizable and volatile variance risk premium consistent with the data. These shocks occur at the late 1990s (LTCM/Russian crisis), the early 2000s (dot.com, 9/11 attacks and the onset of the second Gulf War), and the financial crisis of 2008-2009. The time-varying nature of macroeconomic uncertainty reflects in the variance risk premium, generating short-term predictability for excess returns, consistent with the empirical evidence and previous literature. The model is also able to jointly capture higher order moments of the realized equity premium and salient properties of cash flows and the real risk free rate, with a reasonably low level of relative risk aversion equal to five.

My findings are consistent with some of existing research such as Bollerslev et al. (2009), Drechsler and Yaron (2011), Benzoni et al. (2011) and Drechsler (2013). However, this paper departs from the literature by fully acknowledging parameter and state uncertainty. This feature of the model has several implications. First, belief updates on parameters produce low-frequency shocks in the conditional distribution of consumption growth, generating an endogenous long-run type of risk. This makes consumption growth far less predictable, bridging the gap between ex-ante transitory shocks and their ex-post persistent effect on equilibrium asset prices. Second, parameter uncertainty affects the subjective belief on the state of macroeconomic uncertainty. This enlarges the impact of uncertainty shocks, increasing the variance risk premium in equilibrium. Third, parameter learning reduces the modeler’s degrees of freedom as the structural dynamics is not calibrated but estimated in real-time, with the importance of initial prior information quickly decaying over time (see Martin 2013 and Chen et al. 2013). A more detailed model inspection suggests that parameter uncertainty likely represents a non-diversifiable risk which is priced in equilibrium (see Collin-Dufresne et al. 2013 for a related discussion). In fact, a model with CRRA preferences and parameter learning can account for
a significant fraction of the variance risk premium.

References


Appendix

A Data

A.1 Consumption growth

Consumption data, population and price deflator are from the Federal Reserve Bank of St.Louis. I use personal consumer expenditures on non-durables (PCND) and services consumption (PCESV) deflated by the corresponding chain-type price index (PCEPI). Aggregate consumption is normalized by using total population (POP). The log growth rate of consumption is defined as

$$\Delta c_{t+1} = \ln \left[ \frac{C_{nd,t+1} + C_{s,t+1}}{C_{nd,t} + C_{s,t}} \right]$$  \hspace{1cm} (A.17)

with $C_{s,t}$ and $C_{nd,t}$ the consumption of non-durables and services, respectively. Data are monthly and the sample is 1990:01 - 2013:01.

A.2 Stock market data

The monthly price index for the market is constructed at each time $t$ as (see Campbell and Beeler 2012)

$$P_t = P_{t-1} (1 + VWRETX_t)$$  \hspace{1cm} (A.18)

with VWRETX the value-weighted index excluding share distributions. The price appreciation is adjusted for repurchases as in Bansal et al. (2005),

$$\frac{P_t^*}{P_{t-1}^*} = \left[ \frac{P_t}{P_{t-1}} \right] \min \left[ \left( \frac{n_{t+1}}{n_t} \right), 1 \right]$$

in which $n_t$ represents the number of shares in the market index. The aggregate level of cash dividends is defined as

$$D_t = P_t^* \left[ \frac{(1 + VWRETD_t) - 1}{(1 + VWRETX_t) - 1} \right]$$  \hspace{1cm} (A.19)

with VWRETD the value-weighted index including distributions. Both the aggregate prices and dividends are deflated by the chain-type price index of personal consumption expenditures.

\footnote{As such, $P_t^*/P_{t-1}^*$ represents a downward adjustment of $1 + VWRETX$ only if there is a reduction in the number of shares.}
Gross market index return and the realized equity premium are computed as

\[ R_{t,t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} \]

\[ r_{t,t+1} = \ln (1 + R_{t,t+1}) - r_f^t \]

with \( r_f^t \) the ex-ante real risk free rate. Data are from CRSP and the sample period is 1990:01 - 2013:01, monthly.

### A.3 Inflation

Monthly inflation is the log growth rate of the CPI in the current month over the previous one. Consumption growth is deflated in levels and transformed in log growth rates, while both market returns and the dividend growth are made in real terms by subtracting log inflation from nominal monthly returns.

### A.4 Ex-Ante Risk Free Rate

Nominal yields of the one-month T-Bill rate is from Ibbotson. The ex-post real risk free rate is obtained subtracting the monthly inflation rate \( \pi_{t-1,t} \) to the monthly T-bill yield \( r_{f,t} \) at each time \( t \). The ex-ante real risk free rate \( \hat{r}_{f,t,t+1} \), is computed by projecting the ex-post riskless return \( r_{f,t+1} - \pi_{t,t+1} \) on the average monthly inflation across the previous year \( \pi_{t-12,t} \) and the one month nominal yield \( r_{f,t} \),

\[ r_{f,t,t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 r_{f,t} + \hat{\alpha}_2 \pi_{t-12,t} \]  

(A.20)

### A.5 The Implied and Realized Volatility Measures

The data on the VIX are from the Chicago Board of Options Exchange, i.e. CBOE.\(^{36}\) The measure of realized aggregate returns variation is constructed summing up 78 5-minute squared returns covering a normal trading day, i.e. from 9.30 am to 4.00 pm. For a typical month with 23 trading days I use \( n = 23 \times 78 = 1794 \) squared returns. The high frequency returns are obtained from TICKDATA. The conditional expectation under the physical measure is obtained as \( E_p^t [RV_{t,t+1}] = \alpha + \beta IV_{t-1,t} + \gamma RV_{t-1,t} \), in which \( IV_{t-1,t} = VIX^2_{t-1,t}/12 \) is the VIX index returned in monthly variance terms. Alternative measures investigated are \( E_p^t [RV_{t,t+1}] = RV_{t-1,t} \) and \( E_p^t [RV_{t,t+1}] = \Psi(L) \epsilon_t \) with \( \Psi(L) \) a lag-polynomial of order twelve (Bollerslev et al. 2009).

Let $RV_{t,t+\tau}$ the aggregate market returns variation from time $t$ to $t + \tau$, the variance risk premium is defined as

$$VRP_{t,t+\tau} = E^Q_t [RV_{t,t+\tau}] - E^P_t [RV_{t,t+\tau}]$$ (A.21)

The risk neutral measure can be recovered by using the standard Radon-Nykodim density

$$dQ = \frac{M_{t,t+\tau}}{E^P_t [M_{t,t+\tau}]} dP \quad \text{such that} \quad \frac{dQ}{dP} = \frac{M_{t,t+\tau}}{E^P_t [M_{t,t+\tau}]}$$ (A.22)

in which $M_{t,t+\tau}$ is the stochastic discount factor from $t$ to $t + \tau$. Following Carr and Wu (2009) and Miao et al. (2012) the Variance Risk Premium can be rewritten as

$$VRP_{t,t+\tau} = E^P_t \left[ \frac{M_{t,t+\tau} RV_{t,t+\tau}}{E^P_t [M_{t,t+\tau}]} \right] - E^P_t [RV_{t,t+\tau}]$$

$$= \frac{E^P_t [M_{t,t+\tau} RV_{t,t+\tau}]}{E^P_t [M_{t,t+\tau}]} - E^P_t [RV_{t,t+\tau}]$$ (A.23)

By using the law of iterated expectations

$$E_t [M_{t,t+\tau} RV_{t,t+\tau} | y^t] = E \left[ E_t \left[ M_{t,t+\tau} RV_{t,t+\tau} | z_{t+1}, y^t \right] | y^t \right]$$

$$= \sum_{i=1}^{k} E_t \left[ M_{t,t+\tau} RV_{t,t+\tau} | \mu_{t+1} = i, y^t \right] p \left( \lambda_{t+1} = i | y^t \right)$$

$$= \sum_{i=1}^{k} E^i_{t,\mu} [M_{t,t+\tau} RV_{t,t+\tau}] p \left( \lambda_{t+1} = i | y^t \right)$$

such as

$$E^P_t \left[ RV_{t,t+\tau} | y^t \right] = \frac{\pi^L_{t,t+\tau} E^L_{t,\mu} [M_{t,t+\tau} RV_{t,t+\tau}] + \pi^H_{t,t+\tau} E^H_{t,\mu} [M_{t,t+\tau} RV_{t,t+\tau}]}{\pi^L_{t,t+\tau} E^L_{t,\mu} [M_{t,t+1}] + \pi^H_{t,t+\tau} E^H_{t,\mu} [M_{t,t+\tau}]}$$

in which $y^t$ and $\theta$ the amount of available information and the vector of parameters respectively, $\pi_{t,t+\tau} = p(\lambda_{t+1} = i | y^t)$ the probability of being in the $i$th state one-step ahead, and

$$E^P_{t,\mu} [RV_{t,t+\tau}] = \pi^L_{t,t+\tau} E^L_{t,\mu} [RV_{t,t+\tau}] + \pi^H_{t,t+\tau} E^H_{t,\mu} [RV_{t,t+\tau}]$$

From the definition of conditional covariance the cross-product between the stochastic discount factor and the aggregate returns variation can be rewritten as

$$E_t [RV_{t,t+\tau}, M_{t,t+\tau}] = E_t [RV_{t,t+\tau}] E_t [M_{t,t+\tau}] + Cov_t [RV_{t,t+\tau}, M_{t,t+\tau}]$$
such that
\[
E_t [RV_{t,t+\tau}, M_{t,t+\tau}] = ...
\]
\[
= \frac{\pi_{L,t+\tau}E_{t,\mu}^L [RV_{t,t+\tau}] + \pi_{H,t+\tau}E_{t,\mu}^H [RV_{t,t+\tau}] E_{t,\mu}^H [M_{t,t+\tau}]}{\pi_{L,t+\tau}E_{t,\mu}^L [M_{t,t+\tau}] + \pi_{H,t+\tau}E_{t,\mu}^H [M_{t,t+\tau}]} + \frac{\pi_{L,t+\tau}Cov_{t,\mu}^L [M_{t,t+\tau}, RV_{t,t+\tau}] + \pi_{H,t+\tau}Cov_{t,\mu}^H [M_{t,t+\tau}, RV_{t,t+\tau}]}{\pi_{L,t+\tau}E_{t,\mu}^L [M_{t,t+\tau}] + \pi_{H,t+\tau}E_{t,\mu}^H [M_{t,t+\tau}]} \tag{A.24}
\]
Defining the preferences-adjusted probability of being in a high uncertainty state as
\[
\tilde{\pi}_{H,t+\tau} = \left( \frac{E_{t,\mu}^H [M_{t,t+\tau}]}{\pi_{L,t+\tau}E_{t,\mu}^L [M_{t,t+\tau}] + \pi_{H,t+\tau}E_{t,\mu}^H [M_{t,t+\tau}]} - 1 \right) \pi_{H,t+\tau} \tag{A.25}
\]
and collecting common terms we have
\[
VRP_{t,t+\tau} = \tilde{\pi}_{H,t+\tau} \times (E_{t,\mu}^H [RV_{t,t+\tau}] - E_{t,\mu}^L [RV_{t,t+\tau}]) + ...
\]
\[
+ \frac{(\tau - \tilde{\pi}_{H,t+\tau}) Cov_{t,\mu}^L [RV_{t,t+\tau}, M_{t,t+\tau}]}{E_{t,\mu}^L [M_{t,t+\tau}]} + \frac{\tilde{\pi}_{H,t+\tau} Cov_{t,\mu}^H [RV_{t,t+\tau}, M_{t,t+\tau}]}{E_{t,\mu}^H [M_{t,t+\tau}]} \tag{A.26}
\]
now defining the last covariance terms \( \kappa_t \) the definition in the main text follows.

### C Sequential Bayesian Filtering and Learning

In the model, the agent is uncertain about both the underlying state variables \( z_t = (\lambda_t, \mu_t) \) and the structural parameters \( \theta = (E, \nu, \mu_d, \phi, \sigma_z^2, \sigma_d^2, p_{LL}, p_{HH}, \sigma_L^2, \sigma_H^2) \). The investor holds initial beliefs over the states and parameters \( p(\theta, z_t | y^I) = p(z_t | \theta, y^I) p(\theta | y^I) \) and updates them via sequential Bayes’ rule, in which \( y_t = (\Delta c_t, \Delta d_t) \) and \( y^I = (y_1, ..., y_t) \). The learning scheme occurs in two steps: A prediction step
\[
p(z_{t+1}, \theta | y^I) = \int p(z_{t+1} | z_t, \theta) p(\theta, z_t | y^I) dz_t \tag{A.26}
\]
then belief updating via the predictive likelihood \( p(y_{t+1} | z_{t+1}, \theta) \)
\[
p(z_{t+1}, \theta | y^{t+1}) \propto p(y_{t+1} | z_{t+1}, \theta) p(z_{t+1}, \theta | y^I) \tag{A.27}
\]
which shows the recursive nature of Bayesian updating, as \( p(z_{t+1}, \theta | y^{t+1}) \) is functionally dependent on \( p(z_t, \theta | y^I) \). The main issue is characterize \( p(z_t, \theta | y^I) \) for each time \( t \). Following Carvalho et al. (2010a) and Carvalho et al. (2010b) the joint distribution of states and parameters can be factored out by using a vector of sufficient statistics for the parameters \( \kappa_t \) and the states \( \kappa_t^z \).
such that

\[ p(\theta, \mu_{t+1}, \lambda_{t+1}, \kappa_{t+1}, \kappa_{i,t+1}^x | y^{t+1}) = p(\theta | \kappa_{t+1}) p(\mu_{t+1} | \kappa_{t+1}, \lambda_{t+1}, y^{t+1}) p(\kappa_{t+1}, \kappa_{i,t+1}^x | \lambda_{t+1} | y^{t+1}) \]

In the model \( \kappa_{i,t}^x = (m_t, C_t) \) in which \( m_t \) and \( C_t \) represent the first and second moment of the filtering distribution \( p(\mu_t | y^t, \lambda_t, \kappa_{i,t-1}^x) \), respectively. Given current information \( y_{t+1} \) the uncertainty regime is propagated as

\[ \lambda_{t+1} \sim p(\lambda_{t+1} | \kappa_{i,t}^x, \lambda_t, \theta, \Delta c_{t+1}) \propto p(\Delta c_{t+1} | \mu_t, \lambda_{t+1}) p(\lambda_{t+1} | \lambda_t, \theta) \]

in which \( p(\lambda_{t+1} | \lambda_t, \theta) \) the predictive distribution of the uncertainty state, \( p(\Delta c_{t+1} | \mu_t, \lambda_{t+1}) = N(f_{t+1}, Q_{t+1}) \) the predictive likelihood and

\[
\begin{align*}
 f_{t+1} &= (1 - \nu) E_{\mu} + \nu m_t \\
 Q_{t+1} &= R_{t+1} + \sigma_c^2
\end{align*}
\]

with \( R_{t+1} = \nu^2 C_t + \sigma_{\mu,\lambda_{t+1}}^2 \) the predictive variance of \( \mu_{t+1} \) (see West and Harrison 1997).

Given the current information \( y_{t+1} \) and the simulated uncertainty state \( \lambda_{t+1} \) the drift \( \mu_{t+1} \) is propagated from the conditional sufficient statistics \( \kappa_{t+1} \) and \( \kappa_{i,t+1}^x \)

\[
\begin{align*}
 \kappa_{t+1} &= K(\kappa_t, z_{t+1}, y_{t+1}) \\
 \kappa_{i,t+1}^x &= K^x(\kappa_{i,t}^x, \theta, \lambda_{t+1}, y_{t+1})
\end{align*}
\]

in which \( K^x(\cdot) \) evolves as \( \kappa_{i,t+1}^x = (m_{t+1}, C_{t+1}) \)

\[
\begin{align*}
 m_{t+1} &= a_{t+1} + A_{t+1} e_{t+1} \\
 C_{t+1} &= R_{t+1} - A_{t+1} Q_{t+1}^{-1} A_{t+1}
\end{align*}
\]

with \( a_{t+1} = (1 - \nu) E_{\mu} + \nu m_t \) the predictive mean, \( e_{t+1} = \Delta c_{t+1} - E_{\Delta c_{t+1}} \) the forecasting error and \( A_{t+1} = R_{t+1} Q_{t+1}^{-1} \) the Kalman-gain, such that \( p(\mu_{t+1} | m_{t+1}, C_{t+1}, \lambda_{t+1}, \theta) = N(m_{t+1}, C_{t+1}) \). From the states \( z_{t+1} = (\lambda_{t+1}, \mu_{t+1}) \) and current information \( y_{t+1} = (\Delta c_{t+1}, \Delta d_{t+1}) \), the Kalman-like recursion \( \kappa_{i,t+1} = K(\kappa_t, z_{t+1}, y_{t+1}) \) is updated. The agent’s prior about the drift parameters is a standard normal-inverse gamma prior

\[
P(\beta_\mu | \sigma^2_{\mu,\lambda_t}, \kappa_t) \sim N(b_t, \sigma^2_{\mu,\lambda_t} B_t^{-1})
\]

\[
P(\sigma^2_{\beta_{\mu,t+1}}, \kappa_t) \sim IG(v_{i,t}/2, V_{i,t}/2) \quad i = H, L
\]

in which \( b_t, B_t \) represent respectively a location and scale parameter, \( v_{i,t} \) counts the degrees of freedom for the \( i \)th state and \( V_{i,t} \) is the scale parameter of the inverse-gamma distribution.
After seeing the aggregate consumption at time $t + 1$, prior beliefs are updated as

$$p \left( \beta_{\mu} | \sigma_{c t+1}^2, \kappa_{t+1} \right) \sim N \left( b_{t+1}, \sigma_{\mu, \lambda t+1}^2 B_t^{-1} \right)$$
$$p \left( \sigma_{c t+1}^2 | \kappa_{t+1} \right) \sim IG \left( v_{t+1}/2, V_{t+1}/2 \right)$$

in which the specific Kalman-like recursion evolves as

$$B_{t+1}^{-1} = B_t^{-1} + Z_{t+1} Z_{t+1}'$$
$$B_{t+1}^{-1} b_{t+1} = B_t^{-1} b_t + Z_t \mu_{t+1}$$
$$v_{i, t+1} = v_{i, t} + 1 \text{ for } i = H, L$$
$$V_{i, t+1} = V_{i, t} + \left[ \left( (\mu_{t+1} - Z_t^\top b_{t+1}) \mu_{t+1} + (b_t - b_{t+1}) \right)^\top B_t^{-1} b_t \right] I_{\lambda_{t+1} = i}$$

with $Z_t = [1, \mu_t]$, and $\beta_{\mu} = [(1 - \nu) E_{\mu, \nu}]$. The prior beliefs at time $t$ of the idiosyncratic risk $\sigma_c^2$ is a simple inverse-gamma as

$$p \left( \sigma_c^2 | \kappa_t \right) \sim IG \left( n_t/2, N_t/2 \right)$$

After observing $\Delta c_{t+1}$ and given $\mu_{t+1}$, the posterior is obtained as

$$p \left( \sigma_c^2 | \kappa_{t+1} \right) \sim IG \left( n_{t+1}/2, N_{t+1}/2 \right)$$

in which the scale parameter $N_{t+1}$ and the degrees of freedom $n_{t+1}$ are updated as

$$n_{t+1} = n_t + 1$$
$$N_{t+1} = N_t + (\Delta c_{t+1} - \mu_{t+1})^2$$

As far as the aggregate dividend growth dynamics is concerned, before seeing the data, the agent’s prior is an inverse-gamma and given by

$$p \left( \beta_d | \sigma_d^2, \kappa_t \right) \sim N \left( b_t, \sigma_d^2 B_t^{-1} \right)$$
$$p \left( \sigma_d^2 | \kappa_t \right) \sim IG \left( d_t/2, D_t/2 \right)$$

where $\beta_d = [\mu_d, \phi]$, $B_t$ represents the precision matrix, $b_t$ the sufficient statistics for the conditional mean, $D_t$ the scale parameter for the inverse-gamma density and $d_t$ counts the degrees of freedom. After seeing the aggregate dividend growth at time $t + 1$, conditioned on $\mu_{t+1}$, the posterior beliefs are

$$p \left( \beta_d | \sigma_d^2, \kappa_{t+1} \right) \sim N \left( g_{t+1}, \sigma_d^2 G_{t+1}^{-1} \right)$$
$$p \left( \sigma_d^2 | \kappa_{t+1} \right) \sim IG \left( d_{t+1}/2, D_{t+1}/2 \right)$$
which is updated off-line as

\[ G_{t+1}^{-1} = G_t^{-1} + X_{t+1} X_{t+1}' \]

\[ G_{t+1}^{-1} g_{t+1} = G_t^{-1} g_t + X_t \Delta d_{t+1} \]

\[ d_{t+1} = d_t + 1 \]

\[ D_{t+1} = D_t + \left( \Delta d_{t+1} - X_t^T g_{t+1} \right) \Delta d_{t+1} + (g_t - g_{t+1})^T G_t^{-1} g_t \]

with \( X_t = [1, \mu_{t+1} - E(\mu)]' \). Finally the agent’s beliefs about the transition matrix \( \Pi \) are updated as follows

\[
\begin{align*}
    p_{HH} &\sim \frac{\Gamma(p_{1,t} + p_{2,t})}{\Gamma(p_{1,t}) \Gamma(p_{2,t})} p_{1,t}^{p_{1,t} - 1} (1 - p)^{p_{2,t} - 1} I_{0,1}(p) \\
    p_{LL} &\sim \frac{\Gamma(q_{1,t} + q_{2,t})}{\Gamma(q_{1,t}) \Gamma(q_{2,t})} q_{1,t}^{q_{1,t} - 1} (1 - q)^{q_{2,t} - 1} I_{0,1}(q)
\end{align*}
\]

with

\[
\begin{align*}
    p_{1,t+1} &= p_{1,t} + I(\lambda_{t+1} = H, \lambda_{t} = H) \\
    q_{1,t+1} &= q_{1,t} + I(\lambda_{t+1} = L, \lambda_{t} = L) \\
    p_{2,t+1} &= p_{2,t} + I(\lambda_{t+1} = H, \lambda_{t} = L) \\
    q_{2,t+1} &= q_{2,t} + I(\lambda_{t+1} = L, \lambda_{t} = H)
\end{align*}
\]

the shape parameters of the beta distribution. In words, when both \( \lambda_{t+1} = H \) and \( \lambda_t = H \), the expected probability of being in a high uncertainty state increases. Notice that, under the beta distribution, the expected value of the probability of being in the high uncertainty state at time \( t + 1 \) is computed as

\[
E[p_{HH}|y', \kappa_t] = \frac{p_{1,t+1}}{p_{1,t+1} + p_{2,t+1}}
\]

which is an increasing function of \( p_{1,t+1} \).

### C.1 Hypothesis Testing

Model assessment is done by comparing the marginal posterior models’ probabilities, which are computed integrating out both states and parameter uncertainty.\(^{37}\) Given the specific prior model probability \( p(M_i) \) the posterior probability of model \( i \) is computed as

\[
p(M_i|y') = \frac{p(y'|M_i) p(M_i)}{\sum_{i=1}^{T} p(y'|M_i) p(M_i)}
\]

\(^{37}\)Notice that, by integrating out parameter uncertainty the marginal likelihood punishes needlessly complicated models.
The real-time nature of the learning scheme allows to compute the marginal likelihood recursively as

\[ p(y_t^t | \mathcal{M}_i) = p(y_t^t | y_t^{t-1}, \mathcal{M}_i) p(y_t^{t-1} | \mathcal{M}_i) \]  

(A.30)

where

\[ p(y_t^t | y_t^{t-1}, \mathcal{M}_i) = \int p(y_t^t | y_t^{t-1}, \theta, z_{t-1}, \mathcal{M}_i) p(z_{t-1}, \theta | y_t^{t-1}, \mathcal{M}_i) dz_{t-1} d\theta \]  

(A.31)

The recursive predictive probability

\[ p(y_t^t | y_t^{t-1}, \mathcal{M}_i) \]

is computed from the M particles as

\[ p(y_t^t | y_t^{t-1}, \mathcal{M}_i) \approx \frac{1}{M} \sum_{m=1}^{M} p\left(y_t^t | \left(\theta, \kappa_{t-1}, \kappa_{t-1}^z, z_{t-1}^{(m)}\right), y_t^{t-1}, \mathcal{M}_i\right) \]  

(A.32)

The null hypothesis \( H_0 : \nu = 0 \) against the alternative \( H_1 : \nu \neq 0 \) is tested by using standard Bayes factors. Since the model with \( \nu = 0 \) is nested in the more general unrestricted version the Bayes factors \( BF_{0,1}^t \) might be approximated via the Savage-Dickey density ratio

\[ BF_{0,1}^t = \frac{p(H_0 | y^t)}{p(H_1 | y^t)} = \frac{p(\nu = 0 | y^t, \mathcal{M}_1)}{p(\nu = 0 | y^t, \mathcal{M}_1)} \]  

The denominator can be directly computed from the prior distribution \( p(\nu = 0 | \mathcal{M}_1) \) and the numerator is defined as

\[ p(\nu = 0 | y^t, \mathcal{M}_1) = \int p(\nu = 0 | \theta_{[-\nu]}, \kappa_t, \kappa_t^z) p\left(\theta_{[-\nu]}, \kappa_t, \kappa_t^z | y^t\right) d(\theta_{[-\nu]}, \kappa_t, \kappa_t^z) \]  

(A.33)

with \( \theta_{[-\nu]} \) the vector of parameters without \( \nu \). This posterior can be approximated from the particle weights \( (\theta_{[-\nu]}, \kappa_t, \kappa_t^z)^{(m)} \) as

\[ p(\nu = 0 | y^t, \mathcal{M}_1)^N = \frac{1}{M} \sum_{m=1}^{M} p\left(\nu = 0 | \left(\theta_{[-\nu]}, \kappa_t, \kappa_t^z\right)^{(m)}\right) \]  

(A.34)

Assuming a priori that \( p(H_0) = p(H_1) \), the posterior probability of the null hypothesis can be computed from the Bayes factor \( BF_{0,1}^t \) as

\[ p[H_0 | y^t] = \frac{BF_{0,1}^t}{1 + BF_{0,1}^t} \]  

(A.35)

### D Prior Calibration and Model Assessment

I calibrate the hyper-parameters in the prior distributions both by using a training sample and by referring to the standard consumption-based asset pricing literature. The training sample
consists of the real per-capita consumption growth and real aggregate dividend growth rates from the Robert Shiller’s website.\(^{38}\) The location parameter for \(p(E_\mu)\) is set to be 0.17% on a monthly basis. Likewise the location of \(p(\mu_d)\) is set to be 0.09% on a monthly basis.\(^{39}\) The location of the persistent parameter is set to be \(E[\nu|y^0] = 0.97\), consistent with the long-run risk literature (Bansal and Yaron 2004, Bansal et al. 2007 and Drechsler and Yaron 2011). The prior expected value of the transition probabilities is such that \(P_{HH} = P_{LL} = 0.95\). The location hyper-parameter of the leverage factor is set to be \(E[\phi] = 3.5\) which is in line with Bansal and Yaron (2004), Lettau et al. (2008) and Abel (1999) among the others. Finally the prior mean of \(\sigma^2_c\) is set such that the a priori signal-to-noise ratio is equal to 0.2, which implies a conservative low amount of information brought by the data. The conditional volatility \(\sigma^2_{\mu,t+1}\) is assumed to be higher a priori under \(\lambda_{t+1} = H\) for identification purposes. Table 11 reports the prior hyper-parameters.

I consider a relative low level of confidence on prior information by considering large scales on the prior hyper-parameters. This also ensures a considerable amount of learning through the testing sample. In order to reduce the impact of prior information on final results I cut the first four years of the monthly estimates obtained as a burn-in sample.

As a statistical check on the model ability to effectively fit the data I first report the end-of-sample marginal likelihoods across different model specifications. The marginal likelihood is computed as

\[
p(y^T) = \prod_{t=1}^{T} p(y_t|y_{t-1}) p(y_0) = \frac{1}{N^T} \prod_{t=1}^{T} \sum_{m=1}^{M} p(y_t|(z_t, \theta)^{(i)})
\]

in which \((z_t, \theta)^{(i)}\) are particles from the predictive \(p(z_t, \theta|y^{t-1})\). Table 12 reports the results.

The Bayes factor shows that the model with two regimes and structural learning might be the one preferred by the data. However, the sequential nature of the learning framework also impose to check for the sequential posterior odd probability for the single- vs two-regimes in the dynamics of macroeconomic uncertainty. Figure 13 shows the results. I set the prior model probability \(p(M_i)\) of each model to be 1/2.

The blue area shows the marginal posterior probability of the model with two regimes. The


\(^{39}\)These values correspond to the unconditional means of real per capita annual growth rate of consumption and dividend previous to the 1990 and divided by 12.
single regime model is quickly rejected by the agent that updates belief based on current real per capita consumption.\textsuperscript{40} The null $\mathcal{H}_0 : \nu = 0$ is tested against the alternative $\mathcal{H}_1 : \nu \neq 0$ on a sequential basis by using a Savage-Dickey density ratio. Figure 14 reports the corresponding results across the sample period.

There is evidence of conditional independence from early 2000s, up to the recent great financial crisis. By the end of the sample, however, the single agent learns that the growth rate of real per capita consumption may show significant autocorrelation, although with a low level of predictability.

D.1 Parameters Estimates: End-of-Sample

Here I report the mean, median and 95% percentiles of the end-of-sample parameter estimates of both consumption and dividend dynamics, with $E[E_\mu|y^T]$, for instance, representing the agent’s expectations about the long-run growth rate of consumption once the entire history on fundamentals is available. Table 13 reports the results.

The end-of-sample estimate of the persistence parameter is around 0.2 and statistically significant at the 5% level. The long-run expected growth rate coincides with the historical mean, i.e. $E[E_\mu|y^T] = 0.2110$.

The long-run growth rate of the aggregate dividend, i.e. $E[\mu_d|y^T] = 0.226$ is comparable with the real per capita consumption growth rate on a monthly basis. The leverage parameter is estimated around 2.6, consistent with the previous literature. The high-uncertainty is much less persistent than the low-uncertainty state since $E[P_{HH}|y^T] = 0.512$ as opposed to $E[P_{LL}|y^T] = 0.8688$. Such values imply an average duration of 2 months and around 8 months for high- and low-macroeconomic uncertainty respectively. The unconditional probabilities of the two states are $E[\pi_H|y^T] = 0.22$ and $E[\pi_L|y^T] = 0.78$.\textsuperscript{41} The compounding effect of $\sigma^2_{\mu,\lambda_{t+1}=H}$ and $E[p_{HH}|y^T] = 0.51$ indicates that macroeconomic uncertainty shocks are infrequent but not

\textsuperscript{40}This is consistent with some of the earlier literature such as Cecchetti, Lam, and Mark (1999), Brandt et al. (2004), Guidolin and Timmermann (2007), Ju and Miao (2012) and Cogley and Sargent (2008) just to cite a few.

\textsuperscript{41}The ergodic probability for the $i_{th}$ state is computed as

$$\pi_i = \frac{1 - p_{jj}}{2 - p_{ii} - p_{jj}}$$

while the average duration of the $i_{th}$ state is computed as

$$Dur_i = \frac{1}{1 - p_{ii}}$$
rare, and large but do consistent with the rare disaster literature.

E Numerical Solution of the Model

The dynamics of economic fundamentals is given by

$$\Delta c_{t+1} = \mu_{t+1} + \sigma_c \epsilon_{c,t+1} \quad \epsilon_{c,t+1} \sim N(0, 1)$$

$$\mu_{t+1} = (1 - \nu) E_\mu + \nu \mu_t + \sigma_{\mu,\lambda_{t+1}} \epsilon_{\mu,t+1} \quad \epsilon_{\mu,t+1} \sim N(0, 1)$$

$$\Delta d_{t+1} = \mu_d + \phi (\mu_{t+1} - E_\mu) + \sigma_d \epsilon_{d,t+1} \quad \epsilon_{d,t+1} \sim N(0, 1)$$

The conditional volatility $\sigma_{\mu,\lambda_{t+1}}$ is time-varying and depends on a two-state Markov regime-switching process where the latent regime $\lambda_t = i$, for $i = H, L$ follows the transition probabilities

$$\Pi' = \begin{pmatrix} p_{LL} & 1 - p_{HH} \\ 1 - p_{LL} & p_{HH} \end{pmatrix}$$

The regime changes are assumed to be independent to Gaussian shocks. Both the parameters and the states $z_t = (\lambda_t, \mu_t)$ are assumed to be unknown by the representative agent. The recursive formulation for the wealth-consumption ratio in equilibrium takes the standard form

$$P_t^C = E_t \left[ \beta^\theta \exp^{(1-\gamma) \Delta c_{t+1}} (1 + P_{t+1}^C)^\theta \right]$$

which is the standard formulation under Kreps-Porteus preferences. Notice the conditional expectation $E_t [.]$, is taken with respect to the information available to the agent at time $t$. This information set contains both current information about consumption and dividend growth and the posterior beliefs of the agent on the hidden states and parameters. As such, the optimal wealth-consumption ratio is both belief- and time-dependent. In this setting, unfortunately a closed form solution is simply unfeasible. In the model, there are ten parameters and two unobservable states governing the dynamics of economic fundamentals. The agent’s belief for each of those is governed by two hyper-parameters accordingly, further introducing nuisance parameters, for a total of twenty-four unknowns in the Euler equations. This, makes the state-space prohibitively large, especially considering the optimal policy must be obtained at each time $t$, whatever the sample size could be. Following Kreps 1998, Piazzesi and Schneider 2010, and Cogley and Sargent 2009, I recursively solve the equilibrium model by using an anticipated utility approach. The underlying key implication is that, ex-ante, beliefs on states and parameters are seen as if they will remain constant indefinitely into the future, albeit ex-post they change over time due to learning. This method makes standard numerical methods applicable. Indeed, conditional on the agent’s beliefs both the optimal wealth-consumption ratio and the price-dividend ratio can be found as a solution of a standard fixed-point problem solvable with, for instance,
feasible iterative projection methods. Notice from the dynamics of the conditional sufficient statistics \( \kappa_{t+1} = (m_{t+1}, C_{t+1}) \) that \( m_{t+1} = f(m_t, \Delta c_{t+1}, \lambda_{t+1}, \theta) \), \( \lambda_{t+1} = h(\lambda_t, \Delta c_{t+1}, m_t, \theta) \) and \( \Delta c_{t+1} = g(\kappa_t, \lambda_{t+1}) \). Further, it is necessary to bound the support of the conditional beliefs on the expected growth rate of consumption; \( m_t \in (\overline{m}, \overline{m}) \) and \( C_t \in (\underline{C}, \overline{C}) \). The reason is the same pointed out in Geweke (2001), meaning a non null probability of obtaining extreme values for the conditional expected consumption growth lead to an infinite value of the utility function, violating then the standard trasversaility condition. Given this the wealth-consumption ratio can be rewritten as

\[
P_C(\kappa^x_t, \lambda_t)^\theta = E \left[ \beta^\theta \exp((1-\gamma)\Delta c_{t+1}(\kappa^x_{t+1}, \lambda_{t+1}) (1 + P_C(\kappa^x_{t+1}, \lambda_{t+1}))^\theta \left| \kappa^x_t, \lambda_t \right. \right]
\]

\[
= \beta^\theta E \left[ \exp((1-\gamma)\Delta c_{t+1}(\kappa^x_{t+1}) (1 + P_C(\kappa^x_{t+1}, \lambda_{t+1}))^\theta \left| \kappa^x_t, \lambda_t \right. \right.
\]

\[
= \beta^\theta \sum_{\lambda_{t+1}=L}^H p(\lambda_{t+1}|\lambda_t, \kappa^x_t) \exp((1-\gamma)m_{t+1}^{(1-\gamma)^2}C_{t+1} (1 + P_C(\kappa^x_{t+1}, \lambda_{t+1}))^\theta \left| \kappa^x_t, \lambda_t \right.)
\]

\[
= \beta^\theta \sum_{\lambda_{t+1}=L}^H \kappa^x_{t+1} \exp((1-\gamma)m_{t+1}^{(1-\gamma)^2}C_{t+1} (1 + P_C(\kappa^x_{t+1}, \lambda_{t+1}))^\theta \left| \kappa^x_t, \lambda_t \right.)
\]

where the second equation comes from a straight application of the expectations iteration hypothesis and the least equality from the fact that the uncertainty state \( \lambda_t \) and the exogenous shocks in consumption \( \epsilon_{c,t+1} \) are conditionally indepndent. Notice that the predictive probability of the \( i \)th regime, as well as the conditional moments \( m_{t+1}, C_{t+1} \) are found integrating out parameter uncertainty,

\[
p(\lambda_{t+1}|\lambda_t, \kappa^x_t) = \int p(\lambda_{t+1}|\lambda_t, \theta, \kappa^x_t)p(\lambda_t, \theta|\kappa^x_t) d\theta
\]

where \( \theta \) here indicates the vector of structural parameters. As such, even though the parameters do not directly fit in the optimal policy functions, they still have a huge effect on the equilibrium outcome, especially affecting the time-series dynamics of the growth rate of consumption. Solving for the dividend claim follows a similar argument. By construction the conditional mean and variance of dividend growth still depend on the agent’s beliefs about the expected growth rate of consumption. In fact, under the agent’s filtration

\[
E [\Delta d_{t+1}|t^{t+1}] = \mu_d + E [\mu_{t+1} - E_{\mu}|t^{t+1}] = \mu_d + \nu (m_{t+1} - E_{\mu})
\]

\[
Var [\Delta d_{t+1}|t^{t+1}] = Var [\mu_{t+1} - E_{\mu}|t^{t+1}] + \sigma_d^2 = C_{t+1} + \sigma_d^2
\]
As such the price-dividend ratio can be solved as the fixed point of

\[ P^D (\kappa_t^x, \lambda_t) = E \left[ \beta^\theta \exp \left( -\gamma \Delta \xi_t + \Delta \eta_t + \Delta \xi_{t+1} + \Delta \eta_{t+1} \right) \right] \]

\[ \times P^D (\lambda_{t+1}, \kappa_t^x) \]

Both the optimal wealth-consumption and price-dividend ratios are solved by iterating the above functions until convergence on a grid for \( m_t, C_t \) and considering the state probabilities given an initial guess of \( P^C (\lambda_t, \kappa_t^x) \). The grid is formed by \( G \) points \((m_t^1, ..., m_t^G)\) on the interval \((-5V(m_t), 5V(m_t))\) with \( V(m_t) \) the standard deviation of \( m_t \) at time \( m_t \).

\[ 42 \text{The number of discretizing points } G \text{ is chosen to be odd such that } m_t \text{ is in the middle point of the grid.} \]
Table 1. Descriptive Statistics
This table reports the descriptive statistics for the aggregate market excess returns, the realized and the implied returns variance, the expected value under the physical measure and the corresponding variance risk premium. The sample is 1990:01 - 2013:01, monthly. The table reports the statistics on both a sub sample ending before the recent financial crisis and the full sample. Mkt represents the aggregate market return in excess of the risk free rate. $IV_{t,t+1} = \sqrt{VIX_{t,t+1}^2 / 12}$ is the implied variance measure. $RV_{t,t+1}$ is the realized variance based on high-frequency squared returns. $E_tRV_{t,t+1}$ is computed by $RV_{t,t+1} = \alpha + \beta IV_{t-1,t} + \gamma RV_{t-1,t} + \epsilon_{t+1}$ and $VRP_{t,t+1} = IV_{t,t+1} - E_tRV_{t,t+1}$ the corresponding variance premium.

<table>
<thead>
<tr>
<th></th>
<th>Sub Sample 1990:01-2007:12</th>
<th>Full Sample 1990:01-2013:01</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mkt</td>
<td>IV$_{t,t+1}$</td>
</tr>
<tr>
<td>Mean</td>
<td>0.51</td>
<td>33.48</td>
</tr>
<tr>
<td>Median</td>
<td>1.02</td>
<td>26.74</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>4.06</td>
<td>23.65</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.82</td>
<td>1.98</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.56</td>
<td>9.00</td>
</tr>
<tr>
<td>AR(1)</td>
<td>0.04</td>
<td>0.78</td>
</tr>
<tr>
<td>AR(2)</td>
<td>-0.01</td>
<td>0.60</td>
</tr>
<tr>
<td>min</td>
<td>-17.60</td>
<td>9.04</td>
</tr>
<tr>
<td>max</td>
<td>9.76</td>
<td>163.39</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>47.34</td>
<td>470.49</td>
</tr>
</tbody>
</table>
Table 2. Variance Risk Premium and Long-Run Predictive Regressions

This table reports a set of predictive regressions with $k = 1, \ldots, 24$ months as forecasting horizon. The independent variable is the historical average excess returns over the following $k$ months. The regressor is the historical variance risk premium. The regressions are run with overlapping monthly returns. $SR_{\text{Max}}/SR_{\text{Unc}}$ is the maximum Sharpe ratio attainable under the $i$th model over the unconditional buy-and-hold Sharpe ratio on the aggregate market portfolio. The sample period is 1990:01 - 2013:01. Top panel shows the results of standard Bayesian regressions with Gaussian error terms. Bottom panel shows the results from Bayesian regression with t-distributed errors, which are robust to the impact of outliers. 95% confidence intervals are reported in square brackets.

<table>
<thead>
<tr>
<th>Horizon (Months)</th>
<th>1</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
<th>15</th>
<th>18</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.207</td>
<td>-0.203</td>
<td>-0.004</td>
<td>0.152</td>
<td>0.197</td>
<td>0.233</td>
<td>0.267</td>
<td>0.291</td>
</tr>
<tr>
<td>VRP$_{t,t+1}$</td>
<td>0.347</td>
<td>0.489</td>
<td>0.249</td>
<td>0.167</td>
<td>0.144</td>
<td>0.119</td>
<td>0.097</td>
<td>0.084</td>
</tr>
<tr>
<td>Adj R$^2$ (50%)</td>
<td>2.36</td>
<td>6.82</td>
<td>5.21</td>
<td>3.41</td>
<td>3.20</td>
<td>2.77</td>
<td>2.30</td>
<td>0.82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Horizon (Months)</th>
<th>1</th>
<th>3</th>
<th>6</th>
<th>9</th>
<th>12</th>
<th>15</th>
<th>18</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.333</td>
<td>-0.374</td>
<td>-0.131</td>
<td>0.084</td>
<td>0.116</td>
<td>0.159</td>
<td>0.189</td>
<td>0.211</td>
</tr>
<tr>
<td>VRP$_{t,t+1}$</td>
<td>0.32</td>
<td>0.361</td>
<td>0.252</td>
<td>0.164</td>
<td>0.141</td>
<td>0.126</td>
<td>0.109</td>
<td>0.073</td>
</tr>
<tr>
<td>Adj R$^2$ (50%)</td>
<td>2.21</td>
<td>5.24</td>
<td>4.56</td>
<td>2.94</td>
<td>2.52</td>
<td>2.20</td>
<td>1.84</td>
<td>0.31</td>
</tr>
</tbody>
</table>
Table 3. Variance Risk Premium and Standard Predictors

This table reports a set of one-step-ahead forecasting regressions. The independent variable is the one-month future market excess return. The set of dependent variables contains different measures of the variance risk premium, in addition to a number of standard predictors taken from the literature such as the log price-dividend ratio, the log price-earnings ratio, the Term spread, the Default spread and the Real risk-free interest rate. SR Max/ SR Unc is the maximum Sharpe ratio attainable under the ith model over the unconditional buy-and-hold Sharpe ratio on the aggregate market portfolio. The sample period is 1990:01 - 2013:01. Regression statistics are computed from a robust Bayesian model (corresponding t-stats are in parenthesis).

<table>
<thead>
<tr>
<th>Dependent $r_{m,t+1} - r_{f,t+1}$</th>
<th>Models</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>-0.69</td>
<td>-0.22</td>
<td>-0.34</td>
<td>-0.31</td>
<td>-0.29</td>
<td>-0.02</td>
<td>-0.14</td>
<td>0.14</td>
<td>-0.13</td>
<td>4.90</td>
<td>4.08</td>
<td>11.6**</td>
<td>1.39</td>
<td>13.34***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.52)</td>
<td>(-0.61)</td>
<td>(-0.73)</td>
<td>(-0.57)</td>
<td>(-0.53)</td>
<td>(-0.06)</td>
<td>(-0.51)</td>
<td>(0.38)</td>
<td>(-0.38)</td>
<td>(1.07)</td>
<td>(1.49)</td>
<td>(2.18)</td>
<td>(1.44)</td>
<td>(2.67)</td>
</tr>
<tr>
<td>VRP$_{(M_0)}^{(t,t+1)}$</td>
<td></td>
<td>0.52***</td>
<td>0.57***</td>
<td>0.45***</td>
<td>0.57***</td>
<td>0.45***</td>
<td>0.36***</td>
<td>0.38***</td>
<td>0.48***</td>
<td>0.40***</td>
<td>0.51***</td>
<td>0.33*</td>
<td>0.32**</td>
<td>0.56***</td>
<td>0.60***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3.02)</td>
<td>(3.07)</td>
<td>(2.35)</td>
<td>(4.9)</td>
<td>(3.87)</td>
<td>(3.21)</td>
<td>(2.97)</td>
<td>(2.74)</td>
<td>(1.95)</td>
<td>(2.15)</td>
<td>(2.92)</td>
<td>(3.21)</td>
<td>(2.92)</td>
<td>(3.21)</td>
</tr>
<tr>
<td>IV$_{t,t+1}$</td>
<td></td>
<td>-0.12</td>
<td>0.11</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
<td>-0.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.01)</td>
<td>(0.67)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
<td>(-1.27)</td>
</tr>
<tr>
<td>RV$_{t-1,t}$</td>
<td></td>
<td>-0.07</td>
<td>-0.21</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
<td>-0.28***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.04)</td>
<td>(-1.60)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
<td>(-2.39)</td>
</tr>
<tr>
<td>log($P_t/D_t$)</td>
<td></td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
<td>-1.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
<td>(-1.72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.81)</td>
<td>(-1.93)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
<td>(-2.42)</td>
</tr>
<tr>
<td>TERM$_t$</td>
<td></td>
<td>-0.21</td>
<td>0.22</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.80)</td>
<td>(0.76)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
<td>(-0.48)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.80)</td>
<td>(-2.40)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
<td>(-1.06)</td>
</tr>
<tr>
<td>RealRf$_t$</td>
<td></td>
<td>-1.68</td>
<td>-1.19</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
<td>-1.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.56)</td>
<td>(-0.22)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
<td>(-0.23)</td>
</tr>
<tr>
<td>Adj R$^2$(%)</td>
<td></td>
<td>3.6</td>
<td>3.2</td>
<td>2.2</td>
<td>3.7</td>
<td>3.9</td>
<td>3.4</td>
<td>3.0</td>
<td>2.5</td>
<td>5.4</td>
<td>4.7</td>
<td>6.2</td>
<td>3.8</td>
<td>4.9</td>
<td>7.9</td>
</tr>
<tr>
<td>SR Max/SR Unc</td>
<td></td>
<td>2.3</td>
<td>2.2</td>
<td>1.8</td>
<td>2.3</td>
<td>2.4</td>
<td>2.2</td>
<td>2.1</td>
<td>1.9</td>
<td>2.7</td>
<td>2.5</td>
<td>2.9</td>
<td>2.3</td>
<td>2.6</td>
<td>3.3</td>
</tr>
</tbody>
</table>
Table 4. Variance Risk Premia and Macroeconomic Uncertainty
This table reports a set of regressions in which the I project the variance risk premium on a set of widely used proxies for economic uncertainty. These are the period-by-period cross-sectional dispersion of real consumption growth and real GDP growth, a survey-based market uncertainty index, the predictive variance of consumption growth from a GARCH(1,1) model, the Anxiety index (held by the Philadelphia Fed) and the lagged implied and realized market returns variance. The regressions are estimated through a Bayesian method robust to outliers (t-stats in parenthesis). The sample is 1990:01 - 2013:01, monthly. * stands for statistically significant at 10% confidence level, ** 5% significance and *** statistically significant at the 1% level.

<table>
<thead>
<tr>
<th>Dependent Variable: VRP_{t,t+1}</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>9.06***</td>
<td>2.66</td>
<td>1.02</td>
<td>13.15***</td>
<td>1.06</td>
<td>13.10***</td>
<td>17.43***</td>
</tr>
<tr>
<td></td>
<td>(2.15)</td>
<td>(0.72)</td>
<td>(0.18)</td>
<td>(4.78)</td>
<td>(0.12)</td>
<td>(4.41)</td>
<td>(6.72)</td>
</tr>
<tr>
<td>Cross-Sectional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance of Forecasts on:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real Cons. Growth</td>
<td>21.93**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.33)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real GDP Growth</td>
<td>38.70***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.64)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey-Based Indexes:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Market Uncertainty</td>
<td>0.30***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.71)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxious</td>
<td>0.29***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alternative Measures:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_{t+1</td>
<td>t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50.54**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV_{t-1,t}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.13***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.42)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RV_{t-1,t}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj. R^2</td>
<td>3.6</td>
<td>6.2</td>
<td>16.8</td>
<td>4.2</td>
<td>5.1</td>
<td>4.6</td>
<td>0.57</td>
</tr>
</tbody>
</table>
Table 5. Variance Risk Premium: Unconditional Moments

This table reports the unconditional moments of the historical market variance risk premium, as well as those implied by the model. The preference parameters are \( \gamma = 2.5 \), \( \psi = 3.5 \) and \( \beta = 0.998 \). Panel A and C report the results from the model with real-time structural learning with recursive and CRRA preferences, respectively. Panel B reports the results computed from a rational expectations benchmark, which is obtained by setting the structural parameters of the model equal to the end-of-sample estimates obtained from the model with parameter learning. The sample period is 1990:01 - 2013:01, monthly. \( \mathbb{E}_T \) denotes the ex-post mean computed conditioning on the entire sample path. The first four years of monthly results are removed as a burn-in sample.

### Panel A: Real-Time Structural Learning

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>( \gamma = 5, \psi = 3.5 )</th>
<th>( \gamma = 2, \psi = 3.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (St.Err)</td>
<td>Mean (St.Err) 2.5% 50% 97.5%</td>
<td>Mean (St.Err) 2.5% 50% 97.5%</td>
</tr>
<tr>
<td>VRP</td>
<td>( \mathbb{E}<em>T (\text{VRP}</em>{t+1,t}) )</td>
<td>18.36 (2.44) 17.77 (2.14) 14.27 17.76 21.35</td>
<td>15.72 (1.82) 12.77 15.70 18.79</td>
</tr>
<tr>
<td></td>
<td>( \sigma_T (\text{VRP}_{t+1,t}) )</td>
<td>20.54 (4.57) 19.29 (2.94) 14.19 19.42 23.92</td>
<td>16.48 (2.78) 11.75 16.57 20.96</td>
</tr>
<tr>
<td></td>
<td>skewness</td>
<td>2.16 (0.77) 2.12 (0.45) 1.63 2.17 3.04</td>
<td>2.06 (0.45) 1.41 2.00 2.87</td>
</tr>
<tr>
<td></td>
<td>kurtosis</td>
<td>9.32 (5.60) 9.16 (3.08) 6.11 9.15 15.47</td>
<td>6.88 (3.28) 1.92 5.27 10.91</td>
</tr>
<tr>
<td></td>
<td>( \rho_T (\text{VRP}<em>{t+1,t}, \text{VRP}</em>{t-1,t}) )</td>
<td>0.57 (0.05) 0.53 (0.07) 0.42 0.54 0.64</td>
<td>0.46 (0.06) 0.35 0.46 0.54</td>
</tr>
</tbody>
</table>

### Panel B: Rational Expectations Benchmark

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>( \gamma = 5, \psi = 3.5 )</th>
<th>( \gamma = 2, \psi = 3.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (St.Err)</td>
<td>Mean (St.Err) 2.5% 50% 97.5%</td>
<td>Mean (St.Err) 2.5% 50% 97.5%</td>
</tr>
<tr>
<td>VRP</td>
<td>( \mathbb{E}<em>T (\text{VRP}</em>{t+1,t}) )</td>
<td>18.36 (2.44) 5.84 (0.51) 5.00 5.83 6.68</td>
<td>5.77 (0.50) 4.95 5.76 6.61</td>
</tr>
<tr>
<td></td>
<td>( \sigma_T (\text{VRP}_{t+1,t}) )</td>
<td>20.54 (4.57) 5.41 (0.88) 3.90 5.45 6.79</td>
<td>5.19 (0.91) 3.66 5.20 6.67</td>
</tr>
<tr>
<td></td>
<td>skewness</td>
<td>2.16 (0.77) 2.26 (0.47) 1.62 2.20 4.14</td>
<td>2.23 (0.49) 2.57 3.17 4.11</td>
</tr>
<tr>
<td></td>
<td>kurtosis</td>
<td>9.32 (5.60) 7.70 (2.22) 5.67 7.94 10.63</td>
<td>7.48 (2.18) 4.52 7.69 11.31</td>
</tr>
<tr>
<td></td>
<td>( \rho_T (\text{VRP}<em>{t+1,t}, \text{VRP}</em>{t-1,t}) )</td>
<td>0.57 (0.05) 0.47 (0.07) 0.35 0.48 0.58</td>
<td>0.46 (0.08) 0.31 0.47 0.58</td>
</tr>
</tbody>
</table>

### Panel C: Real-Time Structural Learning (CRRA Preferences)

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>( \gamma = 5 )</th>
<th>( \gamma = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (St.Err)</td>
<td>Mean (St.Err) 2.5% 50% 97.5%</td>
<td>Mean (St.Err) 2.5% 50% 97.5%</td>
</tr>
<tr>
<td>VRP</td>
<td>( \mathbb{E}<em>T (\text{VRP}</em>{t+1,t}) )</td>
<td>18.36 (2.44) 11.48 (1.62) 9.23 11.82 14.61</td>
<td>11.66 (1.90) 8.60 11.64 14.86</td>
</tr>
<tr>
<td></td>
<td>( \sigma_T (\text{VRP}_{t+1,t}) )</td>
<td>20.54 (4.57) 11.20 (2.43) 7.11 11.30 15.06</td>
<td>10.22 (2.72) 5.54 10.27 14.59</td>
</tr>
<tr>
<td></td>
<td>skewness</td>
<td>2.16 (0.77) 1.70 (0.44) 1.11 1.65 2.51</td>
<td>1.67 (0.43) 1.07 1.61 2.46</td>
</tr>
<tr>
<td></td>
<td>kurtosis</td>
<td>9.32 (5.60) 4.35 (1.07) 2.74 4.75 6.01</td>
<td>4.02 (0.97) 2.40 4.48 6.53</td>
</tr>
<tr>
<td></td>
<td>( \rho_T (\text{VRP}<em>{t+1,t}, \text{VRP}</em>{t-1,t}) )</td>
<td>0.57 (0.05) 0.38 (0.06) 0.27 0.39 0.47</td>
<td>0.30 (0.07) 0.17 0.30 0.40</td>
</tr>
</tbody>
</table>
Table 6. Excess Returns Predictability

This table reports the results of a set of model-implied predictive regressions. The dependent variable is the average excess market returns over the following $k$ months, with $k = 1,\ldots,12$. The independent variables are the current model-implied variance risk premium and log price-dividend ratio. Regression estimates are robust for autocorrelation and heteroscedasticity (HAC corrected). The sample is 1990:01 - 2013:01, monthly. The first four years of estimates are removed as a burn-in sample.

**Panel A: Real-Time Structural Learning**

<table>
<thead>
<tr>
<th>Months</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>$\beta^k_{VRP}$ (t-stat)</td>
<td>$\beta^k_{lpd}$ (t-stat)</td>
</tr>
<tr>
<td>1</td>
<td>0.56 (5.37)</td>
<td>-1.77 (-1.72)</td>
</tr>
<tr>
<td>3</td>
<td>0.45 (7.97)</td>
<td>-1.75 (-1.94)</td>
</tr>
<tr>
<td>6</td>
<td>0.28 (5.11)</td>
<td>-2.71 (-2.11)</td>
</tr>
<tr>
<td>9</td>
<td>0.17 (3.04)</td>
<td>-2.95 (-2.48)</td>
</tr>
<tr>
<td>12</td>
<td>0.13 (2.51)</td>
<td>-3.15 (-2.87)</td>
</tr>
</tbody>
</table>

**Panel B: Rational Expectations Benchmark**

<table>
<thead>
<tr>
<th>Months</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>$\beta^k_{VRP}$ (t-stat)</td>
<td>$\beta^k_{lpd}$ (t-stat)</td>
</tr>
<tr>
<td>1</td>
<td>0.56 (5.37)</td>
<td>-1.77 (-1.72)</td>
</tr>
<tr>
<td>3</td>
<td>0.45 (7.97)</td>
<td>-1.75 (-1.94)</td>
</tr>
<tr>
<td>6</td>
<td>0.28 (5.11)</td>
<td>-2.71 (-2.11)</td>
</tr>
<tr>
<td>9</td>
<td>0.17 (3.04)</td>
<td>-2.95 (-2.48)</td>
</tr>
<tr>
<td>12</td>
<td>0.13 (2.51)</td>
<td>-3.15 (-2.87)</td>
</tr>
</tbody>
</table>

**Panel C: Rational Expectations Benchmark ($\nu = 0.98$)**

<table>
<thead>
<tr>
<th>Months</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>$\beta^k_{VRP}$ (t-stat)</td>
<td>$\beta^k_{lpd}$ (t-stat)</td>
</tr>
<tr>
<td>1</td>
<td>0.56 (5.37)</td>
<td>-1.77 (-1.72)</td>
</tr>
<tr>
<td>3</td>
<td>0.45 (7.97)</td>
<td>-1.75 (-1.94)</td>
</tr>
<tr>
<td>6</td>
<td>0.28 (5.11)</td>
<td>-2.71 (-2.11)</td>
</tr>
<tr>
<td>9</td>
<td>0.17 (3.04)</td>
<td>-2.95 (-2.48)</td>
</tr>
<tr>
<td>12</td>
<td>0.13 (2.51)</td>
<td>-3.15 (-2.87)</td>
</tr>
</tbody>
</table>
Table 7. Cash Flows
This table reports the descriptive statistics for the growth rate of real per capita consumption and aggregate dividend growth. The unconditional moments along with standard deviation from non-parametric bootstrap (in parenthesis) are reported in the column Data. Model represent the corresponding unconditional moments and percentiles implied by the model. The sample period is 1990:01 - 2013:01. Cash flows are simulated monthly, then aggregated annually. The first four years of monthly results are removed as a burn-in period.

| Moment          | Data | Model
|-----------------|------|------|
|                 | Estimate (St.Err) | Mean (St.Err) | 2.5% | 50% | 97.5%
| $E_T[\Delta c_t]$ | 2.049 (0.441) | 1.923 (0.271) | 1.592 | 1.919 | 2.372
| $\sigma_T[\Delta c_t]$ | 2.525 (0.405) | 2.368 (0.300) | 1.874 | 2.365 | 2.862
| skew[\Delta c_t] | -0.898 (0.757) | -0.629 (0.206) | -1.073 | -0.623 | -0.108
| kurt[\Delta c_t] | 4.602 (2.104) | 4.736 (0.793) | 3.783 | 4.556 | 5.781
| $\rho(\Delta c_t, \Delta c_{t-1})$ | 0.124 (0.199) | 0.189 (0.206) | 0.098 | 0.176 | 0.281
| $E_T[\Delta d_t]$ | 2.576 (1.794) | 2.234 (1.262) | 0.046 | 2.211 | 4.307
| $\sigma_T[\Delta d_t]$ | 11.521 (1.276) | 10.901 (0.841) | 9.161 | 10.991 | 12.671
| skew[\Delta d_t] | -0.710 (0.819) | -0.618 (0.286) | -1.149 | -0.608 | -0.101
| kurt[\Delta d_t] | 4.222 (0.861) | 3.861 (0.924) | 2.808 | 3.651 | 4.683
| $\rho(\Delta d_t, \Delta d_{t-1})$ | 0.175 (0.192) | 0.171 (0.112) | 0.104 | 0.186 | 0.243
Table 8. Asset Pricing Moments: Equity Returns
This table reports the unconditional moments of aggregate returns, the risk-free rate and the log price-dividend ratio. The preference parameters are $\gamma = 2.5$, $\psi = 3.5$ and $\beta = 0.998$. The returns are computed monthly and aggregated annually. $E_T$ denotes the ex-post mean computed conditioning on the whole history of observables. The sample period is 1990:01 - 2013:01, monthly. The first four years of monthly results are removed as a burn-in sample.

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>$\gamma = 2, \psi = 3.5$</th>
<th>$\gamma = 5, \psi = 3.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (St.Err)</td>
<td>Mean (St.Err) 2.5% 50% 97.5%</td>
<td>Mean (St.Err) 2.5% 50% 97.5%</td>
</tr>
<tr>
<td>Aggregate Equity Premium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_T (R_{mt} - R_{ft})$</td>
<td>4.921 (0.016)</td>
<td>4.201 (1.374) 1.945 4.200 6.469</td>
<td>4.623 (1.402) 2.338 4.611 6.931</td>
</tr>
<tr>
<td>$\sigma_T (R_{mt} - R_{ft})$</td>
<td>17.050 (0.035)</td>
<td>16.328 (0.343) 15.781 16.312 16.906</td>
<td>16.71 (0.367) 16.123 16.704 17.335</td>
</tr>
<tr>
<td>skewness</td>
<td>-0.901 (0.482)</td>
<td>-0.441 (0.115) -0.575 -0.421 -0.318</td>
<td>-0.742 (0.197) -0.749 -0.749 -0.526</td>
</tr>
<tr>
<td>kurtosis</td>
<td>4.072 (1.571)</td>
<td>3.731 (0.837) 2.849 3.542 5.209</td>
<td>3.853 (0.887) 2.912 3.671 5.363</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.335 (0.166)</td>
<td>0.333 (0.111) 0.155 0.331 0.517</td>
<td>0.349 (0.106) 0.173 0.343 0.523</td>
</tr>
<tr>
<td>Log price-dividend ratio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_T (pd_t)$</td>
<td>0.279 (0.042)</td>
<td>0.336 (0.004) 0.311 0.335 0.342</td>
<td>0.421 (0.004) 0.419 0.421 0.423</td>
</tr>
<tr>
<td>$r_T (pd_t, pd_{t-1})$</td>
<td>0.942 (0.016)</td>
<td>0.995 (0.005) 0.985 0.997 0.999</td>
<td>0.949 (0.009) 0.947 0.949 0.950</td>
</tr>
<tr>
<td>Risk-free rate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_T (R_{ft})$</td>
<td>1.021 (0.377)</td>
<td>0.801 (0.133) 0.554 0.805 1.051</td>
<td>0.981 (0.145) 0.661 1.044 1.136</td>
</tr>
<tr>
<td>$\sigma_T (R_{ft})$</td>
<td>1.822 (0.167)</td>
<td>1.861 (0.761) 1.614 2.054 2.746</td>
<td>1.734 (0.031) 1.680 1.731 1.782</td>
</tr>
</tbody>
</table>

Panel A: Real-Time Structural Learning

Panel B: Rational Expectations Benchmark
Table 9. Asset Pricing Moments: Comparative Statics
This table reports a set of comparative statics results. Top panel shows the results from a rational expectations benchmark in which high persistence in the conditional expected growth rate of consumption is exogenously imposed ($\nu = 0.98$). Bottom panel reports the results of model with real-time structural learning and CRRA utility. The preference parameters are $\gamma = 2.5$, $\beta = 0.998$ and $\psi = 3.5$ for the top panel, and $\gamma = 1/\psi$ for power utility. The returns are computed monthly and aggregated annually. $E_T$ denotes the ex-post mean conditioning on the whole history of cash-flows. The sample period is 1990:01 - 2013:01, monthly. The first four years of monthly results are removed as a burn-in sample.

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>$\gamma = 2, \psi = 3.5$</th>
<th>$\gamma = 5, \psi = 3.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (St.Err)</td>
<td>Mean (St.Err)</td>
<td>2.5%</td>
</tr>
<tr>
<td>Aggregate Equity Premium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_T (R_{mt} - R_{ft})$</td>
<td>4.921 (0.016)</td>
<td>2.652 (1.012)</td>
<td>0.625</td>
</tr>
<tr>
<td>$\sigma_T (R_{mt} - R_{ft})$</td>
<td>17.050 (0.035)</td>
<td>14.09 (0.898)</td>
<td>12.53</td>
</tr>
<tr>
<td>skewness</td>
<td>-0.901 (0.482)</td>
<td>-0.219 (0.165)</td>
<td>-0.617</td>
</tr>
<tr>
<td>kurtosis</td>
<td>4.072 (1.571)</td>
<td>2.666 (0.488)</td>
<td>2.032</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.335 (0.166)</td>
<td>0.184 (0.009)</td>
<td>0.169</td>
</tr>
</tbody>
</table>

| Log price-dividend ratio | | | | | | | | | |
| $\sigma_T (pd_t)$ | 0.279 (0.043) | 0.163 (0.027) | 0.123 | 0.161 | 0.212 | 0.359 (0.025) | 0.321 | 0.357 | 0.404 |
| $\rho_T (pd_t, pd_{t-1})$ | 0.942 (0.016) | 0.975 (0.008) | 0.959 | 0.976 | 0.986 | 0.975 (0.008) | 0.961 | 0.976 | 0.986 |

| Risk-free rate | | | | | | | | | |
| $E_T (R_{ft})$ | 1.021 (0.377) | 0.233 (0.107) | 0.032 | 0.226 | 0.826 | 0.887 (0.495) | 0.149 | 0.883 | 1.685 |
| $\sigma_T (R_{ft})$ | 1.822 (0.168) | 0.864 (0.143) | 0.649 | 0.852 | 1.106 | 1.067 (0.167) | 0.810 | 1.054 | 1.369 |

Panel A: Rational Expectations Benchmark ($\nu = 0.98$)

Panel B: Real-Time Structural Learning (CRRA Preferences)

<table>
<thead>
<tr>
<th>Moments</th>
<th>Data</th>
<th>$\gamma = 2$</th>
<th>$\gamma = 5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate (St.Err)</td>
<td>Mean (St.Err)</td>
<td>2.5%</td>
</tr>
<tr>
<td>Aggregate Equity Premium</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$E_T (R_{mt} - R_{ft})$</td>
<td>4.921 (0.016)</td>
<td>1.913 (1.346)</td>
<td>-0.296</td>
</tr>
<tr>
<td>$\sigma_T (R_{mt} - R_{ft})$</td>
<td>17.050 (0.035)</td>
<td>12.718 (0.404)</td>
<td>12.064</td>
</tr>
<tr>
<td>skewness</td>
<td>-0.901 (0.482)</td>
<td>-0.331 (0.192)</td>
<td>-0.511</td>
</tr>
<tr>
<td>kurtosis</td>
<td>4.072 (1.571)</td>
<td>2.987 (0.839)</td>
<td>2.059</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>0.335 (0.166)</td>
<td>0.114 (0.008)</td>
<td>-0.017</td>
</tr>
</tbody>
</table>

| Log price-dividend ratio | | | | | | | | | |
| $\sigma_T (pd_t)$ | 0.279 (0.042) | 0.101 (0.002) | 0.100 | 0.101 | 0.102 | 0.194 (0.002) | 0.192 | 0.194 | 0.196 |
| $\rho_T (pd_t, pd_{t-1})$ | 0.942 (0.016) | 0.884 (0.002) | 0.883 | 0.884 | 0.885 | 0.949 (0.003) | 0.948 | 0.949 | 0.950 |

| Risk-free rate | | | | | | | | | |
| $E_T (R_{ft})$ | 1.021 (0.377) | 0.599 (0.128) | 0.399 | 0.581 | 0.821 | 1.463 (0.117) | 1.232 | 1.442 | 1.844 |
| $\sigma_T (R_{ft})$ | 1.822 (0.167) | 0.033 (0.008) | 0.030 | 0.033 | 0.035 | 0.770 (0.003) | 0.752 | 0.771 | 0.790 |
Table 10. The Role of Real-Time Structural Learning

This table shows the results of a regression of the historical changes in the variance risk premium on the agent’s beliefs updates on the probability of a high macroeconomic uncertainty state. In computing these probabilities, both parameter and states uncertainty have been intergated out. Additional regressors are, belief revisions of both the conditional expected growth rate of consumption and macroeconomic risk, and contemporaneous and lagged consumption growth. The sample period is 1990:01 - 2013:01, monthly. Robust t-stats corrected for Heteroschedasticity and Autocorrelation (HAC) are reported in parenthesis. * stands for statistically significant at 10% confidence level, ** 5% significance and *** statistically significant at the 1% level of confidence. The first four years of estimates are cut as a burn-in sample.

<table>
<thead>
<tr>
<th>Models</th>
<th>Indipendent: VRP$<em>{t,t+1}$ − VRP$</em>{t−1,t}$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>p($\lambda_{t+1}</td>
<td>y^{t+1}$) − p($\lambda_{t+1}</td>
<td>y^{t}$)</td>
<td>0.188** 0.145** 0.159** 0.126* 0.182* 0.132* 0.183* 0.166* 0.141* 0.094*</td>
<td>(2.028) (1.993) (2.001) (1.849) (1.846) (1.799) (1.899) (1.871) (1.892) (1.859)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Beliefs:</td>
<td>E[Δc$_{t+1}</td>
<td>y^{t+1}$] − E[Δc$_{t+1}</td>
<td>y^{t}$]</td>
<td>−0.161* −0.133* −0.097* −0.123** −0.159*</td>
<td>(-0.165) (-1.801) (-1.781) (-0.201) (-1.866)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Var[Δc$_{t+1}</td>
<td>y^{t+1}$] − Var[Δc$_{t+1}</td>
<td>y^{t}$]</td>
<td>0.255** 0.236** 0.224** 0.197*** 0.189*</td>
<td>(1.985) (2.156) (1.925) (2.422) (1.895)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controls:</td>
<td>Δc$_{t+1}$</td>
<td>-0.108 -0.153</td>
<td>(-1.359) (-1.511)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Δc$_{t}$</td>
<td>-0.021 -0.101</td>
<td>(-0.348) (-1.112)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj R$^2$(%)</td>
<td>5.39 6.20 7.71 9.25 9.27 14.8 10.1 10.2 10.3 18.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 11. Prior Calibration
This table reports the hyper-parameters for the model prior calibration. Location represents the location parameter and Scale identifies the scale parameter, for each prior belief. These are calibrated both by using a pre-sample period and by referring to the literature. The calibration sample concerns the pre-1990 annual real per-capita consumption and aggregate dividend growth from the Bob Shiller’s dataset. The scale parameters are set to have prior specifications be uninformative.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Location</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>ν</td>
<td>0.980</td>
<td>0.50</td>
</tr>
<tr>
<td>E_μ</td>
<td>0.170</td>
<td>0.50</td>
</tr>
<tr>
<td>μ_d</td>
<td>0.090</td>
<td>0.50</td>
</tr>
<tr>
<td>φ</td>
<td>3.000</td>
<td>0.50</td>
</tr>
<tr>
<td>P_{LL}</td>
<td>0.950</td>
<td>0.50</td>
</tr>
<tr>
<td>P_{HH}</td>
<td>0.950</td>
<td>0.50</td>
</tr>
<tr>
<td>σ^2_{μ,λ} = H</td>
<td>0.263</td>
<td>0.50</td>
</tr>
<tr>
<td>σ^2_{μ,λ} = L</td>
<td>0.052</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Table 12. Marginal Likelihoods and Bayes Factors
This table reports the end-of-sample marginal likelihoods and the corresponding Bayes factors computed from models with single- vs two-regimes in the conditional volatility of the expected growth rate of consumption. Marginal likelihoods are computed for both the model with real-time parameter learning and the rational expectations benchmark. The Bayes factors are computed by considering the two-regimes with structural uncertainty as the reference model, and > 100 means that the Bayes factor of the benchmark against the corresponding model is greater than 100. The sample is period is 1990:01-2013:01. The first four years of monthly estimates are removed as a burn-in sample.

<table>
<thead>
<tr>
<th></th>
<th>Real-Time Structural Learning</th>
<th>Rational Expectations Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One Regime</td>
<td>Two Regimes</td>
</tr>
<tr>
<td>Marginal Likelihood</td>
<td>-118.301</td>
<td>-107.110</td>
</tr>
<tr>
<td>Bayes Factor</td>
<td>&gt; 100</td>
<td></td>
</tr>
</tbody>
</table>

61
Table 13. Parameters Estimates
This table reports the parameter estimates of the dynamics
\[
\Delta c_{t+1} = \mu_{t+1} + \sigma c_{c,t+1} \\
\Delta d_{t+1} = \mu_d + \phi (\mu_{t+1} - E_\mu) + \sigma d_{d,t+1} \\
\mu_{t+1} = (1 - \nu)E_\mu + \nu\mu_t + \sigma_{\mu,\lambda_t=H}\epsilon_{\mu,t+1} \\
\phi \sim N(0, I_3)
\]
where \(\lambda_t = i\), for \(i = H, L\) follows a transition probability
\[
p(\lambda_{t+1} = H|\lambda_t = H, \theta) = p_{HH} \quad \text{and} \quad p(\lambda_{t+1} = L|\lambda_t = L, \theta) = p_{LL}
\]
I report the model estimates at time \(T\) corresponding to a MLE-like full-sample estimates. \(E_T\) denotes the ex-post mean computed conditioning on the whole history of cash-flows. The sample is, 1990:01 to 2013:01. The first four years of monthly estimates are removed as a burn-in sample.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>(St.Err)</th>
<th>2.5%</th>
<th>50%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\nu)</td>
<td>0.2081</td>
<td>(0.0571)</td>
<td>0.1206</td>
<td>0.2007</td>
<td>0.3022</td>
</tr>
<tr>
<td>(E_\mu)</td>
<td>0.1910</td>
<td>(0.0212)</td>
<td>0.1710</td>
<td>0.2110</td>
<td>0.2320</td>
</tr>
<tr>
<td>(\sigma^2_{\mu,\lambda_t=L})</td>
<td>0.0321</td>
<td>(0.0036)</td>
<td>0.0310</td>
<td>0.0331</td>
<td>0.0353</td>
</tr>
<tr>
<td>(\sigma^2_{\mu,\lambda_t=H})</td>
<td>0.2416</td>
<td>(0.0732)</td>
<td>0.1507</td>
<td>0.2279</td>
<td>0.3820</td>
</tr>
<tr>
<td>(\mu_d)</td>
<td>0.2265</td>
<td>(0.0967)</td>
<td>0.0807</td>
<td>0.2267</td>
<td>0.3816</td>
</tr>
<tr>
<td>(\phi)</td>
<td>2.6212</td>
<td>(0.7762)</td>
<td>1.3492</td>
<td>2.6321</td>
<td>3.9012</td>
</tr>
<tr>
<td>(p_{HH})</td>
<td>0.5129</td>
<td>(0.0005)</td>
<td>0.5121</td>
<td>0.5129</td>
<td>0.5136</td>
</tr>
<tr>
<td>(p_{LL})</td>
<td>0.8688</td>
<td>(0.0002)</td>
<td>0.8685</td>
<td>0.8688</td>
<td>0.8690</td>
</tr>
<tr>
<td>(\pi_L)</td>
<td>0.7747</td>
<td>(0.0003)</td>
<td>0.7743</td>
<td>0.7747</td>
<td>0.7751</td>
</tr>
<tr>
<td>(\pi_H)</td>
<td>0.2253</td>
<td>(0.0004)</td>
<td>0.2257</td>
<td>0.2253</td>
<td>0.2249</td>
</tr>
<tr>
<td>(Dur_L)</td>
<td>7.0577</td>
<td>(0.0010)</td>
<td>7.0313</td>
<td>7.0578</td>
<td>7.0839</td>
</tr>
<tr>
<td>(Dur_H)</td>
<td>2.0530</td>
<td>(0.0012)</td>
<td>2.0496</td>
<td>2.0530</td>
<td>2.0559</td>
</tr>
</tbody>
</table>

Panel B: Conditional Belief

<table>
<thead>
<tr>
<th>Mean</th>
<th>(St.Err)</th>
<th>2.5%</th>
<th>50%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(E_t[\mu_{t+1}])</td>
<td>0.2151</td>
<td>(0.0113)</td>
<td>-0.3639</td>
<td>0.2300</td>
</tr>
<tr>
<td>(Std_t[\mu_{t+1}])</td>
<td>0.0312</td>
<td>(0.0160)</td>
<td>0.0135</td>
<td>0.0261</td>
</tr>
</tbody>
</table>
Figure 1. Variance Risk Premium and Macroeconomic Uncertainty
This figure shows the market variance risk premium against three alternative measures of macroeconomic uncertainty. The sample period is 1990:01 - 2013:01, monthly
Figure 2. Conditional Variance Real Consumption Growth
This figure shows the squared residuals from an AR(12) model fitted on real consumption growth. The sample period is 1990:01 - 2013:01, monthly.
Figure 3. Impulse Responses
This figure shows the impact of a shock in different proxies of macroeconomic uncertainty on the historical variance risk premium. Top panel shows the impact of a one-standard deviation shock in the variance of real consumption growth forecasts (one-step ahead). Mid panel shows the impact of a one-standard deviation shock in the Anxious index. Bottom panel shows the impact of a one-standard deviation shock in a survey-based measure of market uncertainty. Impulse responses are computed from a VAR(1) model.
Figure 4. Real-Time Parameters Estimates
This figure shows the time series of the posterior estimates of the persistence parameter (top panel) and the unconditional growth rate of consumption (bottom panel). The darker line represents the median value and the gray area plots the posterior distribution at each time $t$. The sample period is 1990:01 - 2013:01. The first four years of monthly estimates are cut as a burn-in sample.
Figure 5. Uncertainty Shocks
Panel A shows the time-varying probability of a high uncertainty state. The sample period is 1990:01 - 2013:01, monthly. The first four years of monthly estimates are cut as a burn-in sample.

Figure 6. Matching the Conditional Dynamics
This figure shows the historical (red line) and the model-implied (blue line) conditional dynamics of the variance risk premium. The sample period is 1990:01 - 2013:01, monthly. The first four years of monthly estimates are cut as a burn-in sample.
Figure 7. Agent’s Belief on the Expected Growth Rate of Consumption
Panel A shows the unconditional distribution of the agent’s belief on the expected growth rate of consumption. The blue line represents the subjective belief, while the red dashed line is a Gaussian benchmark. Panel B shows the behavior of the belief’s revision across the probability of a state of high macroeconomic uncertainty.
Figure 8. The Impact of Uncertainty Shocks
This figure shows the relation between the state $\lambda_{t+1}$ on both the model-implied level of macroeconomic uncertainty (left scale), and the predictive variance of the consumption computed from a GARCH(1,1) model (right scale).
Figure 9. Posterior Belief Dispersion and the Variance Risk Premium
This figure shows the relation between the agent’s posterior belief and the market variance risk premium. Top panel shows the dispersion of the agent’s belief about the expected growth rate of consumption against the variance risk premium. Bottom panel shows the relation between the predictive variance of consumption growth from a GARCH(1,1) model and the variance risk premium. The dashed black line represents the slope of a least squares projection. The sample period is 1990:01 - 2013:01, monthly.
Figure 10. The Role of Parameter Learning
This figure shows the risk-adjusted \( \tilde{\pi}_{H,t+1} \) as a function of the predictive probability \( \pi_{H,t+1} \). This probability is computed from both the model with real-time structural learning and the rational expectations benchmark. The stochastic discount factor is obtained with \( \gamma = 5 \) and \( \psi = 3.5 \).

Figure 11. Uncertainty State Probability: Structural Learning vs Rational Expectations
This figure shows the difference between the predictive probabilities of a high uncertainty state computed from the model with real-time structural learning and the rational expectations benchmark (end-of-sample calibration). The sample period is 1990:01 - 2013:01, monthly. The first four years of monthly estimates are cut as a burn-in sample.
Figure 12. The Role of Risk Aversion
This figure shows the behavior of the risk-adjusted $\tilde{\pi}_{H,t+1}$ as a function of the level of risk aversion. The stochastic discount factor is computed under $\gamma = 2.5$ and $\psi = 3.5$.

Figure 13. Posterior Models Probabilities
This figure shows sequential posterior model probabilities of the two-states model vs. the single state dynamics. The model posterior probabilities are computed from the model with real-time structural learning. The sample period is 1990:01 - 2013:01, monthly. The first four years of monthly estimates are cut as a burn-in sample.
Figure 14. Hypothesis Testing
This figure reports the $p(H_0|y')$ where the null hypothesis $H_0 : \nu = 0$ and $H_1 : \nu \neq 0$. The sample period is 1990:01 - 2013:01, monthly. The first four years are cut as burn-in sample.