PhD THESIS DECLARATION

The undersigned

SURNAME	Bianchi
FIRST NAME	Daniele
PhD Registration Number	1369917

Thesis title:

Essays in Asset Pricing

PhD in	Finance
Cycle	XXIV
Candidate's tutor	Prof. Massimo Guidolin
Year of thesis defence	2014

DECLARES

Under his/her responsibility:

1) that, according to Italian Republic Presidential Decree no. 445, 28th December 2000, mendacious declarations, falsifying records and the use of false records are punishable under the Italian penal code and related special laws. Should any of the above prove true, all benefits included in this declaration and those of the temporary embargo are automatically forfeited from the beginning;

2) that the University has the obligation, according to art. 6, par. 11, Ministerial Decree no. 224, 30th April 1999, to keep a copy of the thesis on deposit at the "Bilioteche Nazionali Centrali" (Italian National Libraries) in Rome and Florence, where consultation will be permitted, unless there is a temporary embargo protecting the rights of external bodies and the industrial/commercial exploitation of the thesis;

3) that the Bocconi Library will file the thesis in its "Archivio istituzionale ad accesso aperto" (institutional registry) which permits online consultation of the complete text (except in cases of a temporary embargo);

4) that, in order to file the thesis at the Bocconi Library, the University requires that the thesis be submitted online by the candidate in unalterable format to Società NORMADEC (acting on behalf of the University), and that NORMADEC will indicate in each footnote the following information:

- thesis (*thesis* title) ....Essays in Asset Pricing...
.... ;

- by (*candidate's surname and first name*) ....Bianchi Daniele........ ;
- defended at Università Commerciale "Luigi Bocconi" – Milano in (year of defence) ..2014........ ;
- the thesis is protected by the regulations governing copyright (Italian law no. 633, 22nd April 1941 and subsequent modifications). The exception is the right of Università Commerciale "Luigi Bocconi" to reproduce the same for research and teaching purposes, quoting the source;

- **only in cases where another declaration has been undersigned requesting a temporary embargo**: the thesis is subject to a temporary embargo for (indicate duration of the embargo) .... months;

5) that the copy of the thesis submitted online to NORMADEC is identical to the copies handed in/sent to the members of the Thesis Board and to any other paper or digital copy deposited at the University offices, and, as a consequence, the University is absolved from any responsibility regarding errors, inaccuracy or omissions in the contents of the thesis;

6) that the contents and organization of the thesis is an original work carried out by the undersigned and does not in any way compromise the rights of third parties (Italian law, no. 633, 22nd April 1941 and subsequent integrations and modifications), including those regarding security of personal details; therefore the University is in any case absolved from any responsibility whatsoever, civil, administrative or penal, and shall be exempt from any requests or claims from third parties;

7) **Choose between 7a or 7b below**:

7a) that the PhD thesis is not the result of work included in the regulations governing industrial property, was not produced as part of projects financed by public or private bodies with restrictions on the diffusion of the results, and is not subject to patent or protection registrations, and therefore not subject to an embargo;

Date __25/01/2014______

SURNAME Bianchi

FIRST NAME Daniele

Essays in Asset Pricing

Daniele Bianchi

daniele.bianchi@phd.unibocconi.it

Thesis submitted to

Università Bocconi

in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Finance

Department of Finance
Via Röentgen 1
I-20136 Milano, Italy

Submitted on January 15, 2014

for Myself

Acknowledgements

There are undoubtedly other friends and colleagues who I have forgotten to thank. If you feel unjustly left out of the list, I apologize. Please go ahead and add yourself here: [ ].

Contents

Preface

In Chapter 1, I show how time-varying macroeconomic uncertainty can reconcile the puzzling disconnect between the macroeconoy and the market premium for the variance risk. By using structural learning within a single-agent endowment economy, I show that economic uncertainty, which means uncertainty on the dynamics of economic fundamentals, may effectively help explain some of the key features of the market premium for the variance risk such as its higher-order unconditional moments, the short-term predictive power on the market excess return and the conditional dynamics. In doing so I relax some of the crucial assumptions which are usually imposed in the reference literature such as full-information about the structural parameters of the economy, jump-like shocks in macroeconomic risk as well as I can get rid of the exogenously persistence which is usually imposed in standard approaches such as the extended long-run risk framwework of Drechsler and Yaron (2011) and Benzoni, Collin-Dufresne, and Goldstein (2011) among the others. I shows, indeed, that I can generate sensible option-pricing implications even with a conditionally i.i.d consumption growth under kreps-porteus preferences. As such, I provide evidence that structural uncertainty may represents an extra-source of non-diversifiable risk which is heavily priced in equilibrium, having therefore a relevant effect on the dynamics of the pricing kernel, above and beyond what may be justified with risk aversion and investor's preferences.

Chapter 2 is co-authored with Massimo Guidolin and Francesco Ravazzolo. This chapter aims at solving the puzzling empirical evidence that a small set of macroeconomic risk factors can indeed explain the cross-sectional variation of stock and bond returns in the U.S. This apparent disconnect has long been representing a puzzle since clashes with a large body of theoretical literature which argues that, indeed, macroeconomic risk factors (i.e. systematic risk) may explain the cross-section of stock and bond returns for the most part. We propose and estimate a novel linear multi-factor asset pricing model which flexibly allows for time variation in both the macroeconomic risk exposures, idiosyncratic volatility and risk premia. We find that, in fact, a small set of macroeconomic risk factors, such as market risk, real output growth, inflation risk and liquidity can explain upto 80% of the cross-sectional variation of stock and bond returns in the U.S. On the asset pricing side this reflects in a better understanding of how macroeconomic factors can explain the dynamics of the pricing kernel. This, of course, is not only crucial for asset and derivatives pricing but also for capital budgeting decisions. Above and beyond the asset pricing implications, we also make a methodological contribution. The model proposed, indeed, allows to get rid of the main drawbacks which typically affects standard, still becnchmarking, two-steps naive-type rolling window regression models such as the so-called Fama-MacBeth cross-sectional regression setting. We test the model performances and show that a model which allows for random breaks in the dynamics of the structural parameters is favored by the data. This is true not only against the aforementioned two-steps procedure, but also with respect to a plethora of dynamic models such as the one proposed by Jostova and Philipov (2005), for instance.

The final chapter concerns about the linkages between the dynamics of macroeconomic fundamentals and the higher-moments of market excess returns. By considering a similar endowment economy with incomplete information as proposed in Chapter 1, I show that time-varying economic uncertainty may indeed help explain higher-order moments of the market excess returns. Again, this is done relaxing some of the usual crucial assumptions such as full-information, long-run risk and jump-like shocks in conditional volatility. Notice that even though the modelling framework is relatively similar to the one proposed in Chapter 1, both the research question, the data and the model implications are fairly different. On top of that in this chapter I propose a novel framework which allows to sequentially learn about state variables and structural parameters in a real-time fashion. This not only helps to get rid of the long-run risk channel, as recently proposed by Collin-Dufresne, Johannes, and Lochstoer (2013), but also allows to flexibly consider the comprehensive amount of uncertainty the marginal investor must face once has been endowed with the same amount of information availble to the econometrician.

Chapter 1

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

JEL codes: G12, G13, E44, C11

1.1 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 C.6 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

[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.

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 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 chapter, I show that time-varying uncertainty on economic fundamentals generate a large and volatile variance risk premium. Uncertainty shocks are large, 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 manifests in a more pessimistic pricing kernel, namely increasing the discount rates.[3] 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

[3]Simply 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.

Bollerslev, Tauchen, and Zhou (2009). Finally, unlike the reference literature, the model can help replicate the dynamics of the variance risk premium. Notably this dynamics is generated endogenously without directly using information available from financial markets.

The framework designed in this chapter includes recursive Kreps-Porteus preferences and real-time learning on state variables and structural parameters.[4] 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.[5] 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 et al. 2013 for 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.

[4]More details on recursive preferences are in Kreps and Porteus (1978), Epstein and Zin (1989), Epstein and Zin (1991) and Weil (1989).

[5]These parameters are those the reference literature assume to be observable by the agent.

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.[6] At a more general level, these results, may help reach a better understanding of the economics behind the large and volatile premium for the variance risk, namely the variance risk premium.

This chapter fits into a growing literature that aims to solve the so-called variance risk premium puzzle, namely the apparent disconnect between the dynamics of economic fundamentals and the 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,

[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.

Distaso, and Mele (2013), reveal that the variance risk premium is strongly negatively related to the business cycle. The equilibrium model in this chapter incorporate these characteristics and is consistent with these findings.

Despite this large amount of empirical evidence, few works 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 et al. (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).[7] 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. 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 chapter, 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 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

[7] 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."*

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.

Methodologically, this chapter represents one of the first attempts, together with Johannes, Lochstoer, and Mou (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.[8] Simultaneous learning about multiple unknowns is able to better capture uncertainty on the dynamics of consumption-based asset pricing models. In this chapter, 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 conditional dynamics of consumption growth, still preserving the standard assumption of conditionally Gaussian innovations in the data generating process.

[8]See for instance Barsky and De Long (1993), Timmermann (1993), Timmermann (1996), Veronesi (1999), Veronesi (2000), Brennan and Xia (2002), Lewellen and Shanken (2002), Brandt, Zeng, and Zhang (2004), Massa and Simonov (2005), Guidolin and Timmermann (2007), Hansen (2007), Weitzman (2007),Lettau, Ludvigson, and Watcher (2008), Hansen et al. (2008), Bansal and Shaliastovich (2011), Bakshi and Skoulakis (2010) and Bansal and Shaliastovich (2011) among the others.

1.2 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 τ;

$$VRP_{t,t+\tau} = E_t^{\mathbb{Q}}\left[RV_{t,t+\tau}\right] - E_t^{\mathbb{P}}\left[RV_{t,t+\tau}\right] \tag{1.1}$$

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, $E_t^{\mathbb{P}}[\cdot]$ and $E_t^{\mathbb{Q}}[\cdot]$ 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

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

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).[9]

Although a high-frequency discretized measure of the realized variance is widely accepted in the literature, the method for constructing the physical expectation $E_t^{\mathbb{P}}\left[RV_{t,t+1}\right]$ 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).[10] 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.[11]

Market returns correspond to the value-weighted return of NYSE,AMEX and NAS-

[9]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.

[10]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).

[11]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 chapter does not add anything to this debate.

DAQ, 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 C.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.[12] As the table shows, the variance risk premium is not normally distributed, in fact, a standard Jarque-Bera

[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.

test strongly rejects the null of $VRP_{t,t+1}$ be Gaussian distributed.

Table C.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.[13] The table reports posterior median estimates, as well as the confidence 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.[14] 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.[15] Table C.6 reports the regression results from robust Bayesian regressions.

[13] More details on the Bayesian regressions can be found in a separate online appendix.

[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 $\left(SR^{\mathcal{H}}\right)^2 = SR_0^2 + \frac{1+SR_0^2}{1-R^2}R^2$ where $SR^{\mathcal{H}}$ the maximum unconditional Sharpe ratio under the model $\mathcal{H}$, 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

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.

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 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).

1.2.1 Variance Risk Premium and Macroeconomic Uncertainty

At the outset of the chapter, I argue that there may be an economically plausible relation between macroeconomic uncertainty and the variance risk premium. Figure C.6 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,

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.

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 B.6 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 B.6. Figure C.1 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 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.

[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.

1.3 The Asset Pricing Model

1.3.1 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

$$V_t = \left\{ (1-\beta) C_t^{1-\frac{1}{\psi}} + \beta \mathcal{R}_t \left(V_{t+1}\right)^{1-\frac{1}{\psi}} \right\}^{\frac{1}{1-\psi}} \tag{1.2}$$

in which $\psi \neq 0$ is the coefficient of intertemporal elasticity of substitution (IES), $\gamma \neq 1$ the relative risk aversion (RRA), and β the subjective discount factor. When the relative risk aversion exceeds the inverse of the IES, the investor prefers early resolution of uncertainty. Let P_t^C 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

$$M_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\beta \frac{1+P_{t+1}^C}{P_t^C}\right)^{\frac{1/\psi-\gamma}{1-1/\psi}} \tag{1.3}$$

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

$$W_{t+1} = W_t \left(1 - k_t\right) R_{c,t+1}$$

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

$$E\left[M_{t,t+1}R_{c,t+1}|y^t\right] = 1, \qquad R_{c,t+1} = \frac{P^C_{t+1} + C_{t+1}}{P^C_t} \tag{1.4}$$

$$E\left[M_{t,t+1}R_{d,t+1}|y^t\right] = 1, \qquad R_{d,t+1} = \frac{P^D_{t+1} + D_{t+1}}{P^D_t} \tag{1.5}$$

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

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 $\log(R_{d,t+1}) - r_{f,t+1}$. 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.

[17] Campbell and Beeler (2012) estimated a value of ψ 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.

1.3.2 Dynamics of Economic Fundamentals

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

$$\Delta c_{t+1} = \mu_{t+1} + \sigma_c \epsilon_{c,t+1} \qquad \epsilon_{c,t+1} \sim N(0,1) \tag{1.6}$$

with σ_c the idiosyncratic volatility. The expected growth rate of consumption μ_{t+1} is a latent AR(1) process with ν the level of persistence of the Gaussian and stationary innovation $\epsilon_{\mu,t+1}$.

$$\mu_{t+1} = (1-\nu)E_\mu + \nu\mu_t + \sigma_{\mu,\lambda_{t+1}}\epsilon_{\mu,t+1} \qquad \epsilon_{\mu,t+1} \sim N(0,1) \tag{1.7}$$

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

$$\Pi' = \begin{pmatrix} p_{LL} & 1-p_{HH} \\ 1-p_{LL} & p_{HH} \end{pmatrix} \tag{1.8}$$

in which

$$p\left(\lambda_{t+1} = H|\lambda_t = H, \theta\right) = p_{HH} \qquad \text{and} \qquad p\left(\lambda_{t+1} = L|\lambda_t = L, \theta\right) = p_{LL} \tag{1.9}$$

and $\sigma^2_{\mu,\lambda_{t+1}} \in \left\{\sigma^2_{\mu,H}, \sigma^2_{\mu,L}\right\}$ with $\sigma^2_{\mu,H} > \sigma^2_{\mu,L}$. 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 λ_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 Δd_{t+1} is defined as a leveraged version of the growth rate of consumption with $\phi > 1$ the rescaling factor.

$$\Delta d_{t+1} = \mu_d + \phi\left(\mu_{t+1} - E_\mu\right) + \sigma_d \epsilon_{d,t+1} \tag{1.10}$$

Notice, inference on the long-run expected growth rate of both consumption and dividend are not influenced by λ_{t+1}. In fact, the unconditional mean of consumption growth collapses to E_μ, such as $E\left[E_\mu | y^T\right]$ coincides with the true unconditional mean as $T \to \infty$. Since $E\left[\mu_{t+1} - E_\mu\right] = 0$, also the μ_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. $[\epsilon_{c,t+1}, \epsilon_{\mu,t+1}, \epsilon_{d,t+1}]' \sim N(0, I_3)$.

1.3.3 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 $\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{M_{t,t+1}}{E_t^{\mathbb{P}}[M_{t,t+1}]}$ 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[18]

$$\begin{aligned} VRP_{t,t+1} &= E_t^{\mathbb{P}}\left[\frac{M_{t,t+1}}{E_t^{\mathbb{P}}\left[M_{t,t+1}\right]} RV_{t,t+1}\right] - E_t^{\mathbb{P}}\left[RV_{t,t+1}\right] \\ &= \frac{E_t^{\mathbb{P}}\left[M_{t,t+1} RV_{t,t+1}\right]}{E_t^{\mathbb{P}}\left[M_{t,t+1}\right]} - E_t^{\mathbb{P}}\left[RV_{t,t+1}\right] \end{aligned} \tag{1.11}$$

in which $E_t^{\mathbb{P}}\left[M_{t,t+1}\right]$ 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,

[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.

the variance risk premium is calculated as

$$VRP_{t,t+1} = \tilde{\pi}_{H,t+1} \times \left(E^H_{t,\mu}\left[RV_{t,t+1}\right] - E^L_{t,\mu}\left[RV_{t,t+1}\right] \right) + \kappa_t \tag{1.12}$$

The first component in (1.12) is convex in $p(\lambda_{t+1} = H|y^t)$, and bounded at zero when either $\pi^{(H)}_{t+1|t} = 0$ 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

$$\tilde{\pi}_{H,t+1} = \left(\frac{E^H_{t,\mu}\left[M_{t,t+1}\right]}{\pi_{L,t+1} E^L_{t,\mu}\left[M_{t,t+1}\right] + \pi_{H,t+1} E^H_{t,\mu}\left[M_{t,t+1}\right]} - 1 \right) \pi_{H,t+1} \tag{1.13}$$

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, $E^i_{t,\mu}\left[M_{t,t+1}\right]$ represents the expected value of the SDF under recursive preferences at time t, given the ith uncertainty regime and the drift μ_{t+1}. The second component is always positive since $Var_{t,H}\left[\mu_{t+1}\right] > Var_{t,L}\left[\mu_{t+1}\right]$ generates $Var^H_{t,\mu}\left[RV_{t,t+1}\right] > Var^L_{t,\mu}\left[RV_{t,t+1}\right]$ under recursive preferences for early resolution of uncertainty. The third component κ_t is convex in $\tilde{\pi}_{H,t+1}$ and is positive since $Cov^i_{t,\mu}\left[RV_{t,t+1}, M_{t,t+1}\right] > 0$ although numerically negligible (see the appendix). I leave more complete analytical expressions to the appendix.

[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.

1.3.4 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 chapter 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 $\theta = (E_\mu, \nu, \mu_d, \phi, \sigma_c^2, \sigma_d^2, p_{LL}, p_{HH}, \sigma_L, \sigma_H)$, 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

[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 constrast, 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. 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."*

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_μ is higher than previously thought, the belief revision mechanism generates a permanent 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 $p\left(\theta, z_t|y^t\right) = p\left(z_t|\theta, y^t\right)p\left(\theta|y^t\right)$ 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

$$p(z_{t+1}, \theta|y^t) = \int p(z_{t+1}|z_t, \theta)p(z_t, \theta|y^t)dz_t \tag{1.14}$$

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

$$p(z_{t+1}, \theta|y^{t+1}) \propto p(y_{t+1}|z_{t+1}, \theta)p(z_{t+1}, \theta|y^t) \tag{1.15}$$

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 λ_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). This method not only keeps the computational burden feasible, but also allows to estimate the model dynamics such as the equilibrium model is solved from a purely forward looking perspective as implied by the Euler equations (1.5). I leave to the appendix a more detailed explanation of the computational strategy.

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

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

with $E\left[\mu_{t+1}|y^t\right]$ computed from the predictive distribution $p\left(\mu_{t+1}|y^t\right)$. This predictive is obtained by integrating out uncertainty on both parameters and the state λ_{t+1}, which makes it heavy-tailed, then leading the agent to see the economic outlook as riskier.[22] Under the agent's filtration, macroeconomic uncertainty is defined as

$$Var\left[\mu_{t+1}|y^t\right] = \nu^2 C_t + \sigma_\mu^2$$

where $C_t = Var\left[\mu_t|y^t\right]$ the posterior variance of μ_t, ν the persistence parameter and σ_μ^2 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 ν, 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

[22] In a recent chapter 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.

of consumption conditional on the state λ_{t+1},

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

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

$$Var\left[\mu_{t+1}|y^t\right]=\sum_{i=1}^{k}p(\lambda_{t+1}=i|y^t)\times Var\left[\mu_{t+1}|\lambda_{t+1}=i,y^t\right] \tag{1.16}$$

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.

1.4 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.

1.4.1 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

$$E\left[\theta|y^t\right] = \int \theta p\left(\theta, z_t|y^t\right) d\theta dz_t$$

with $y^t = (y_1, ..., y_t)$ the available information and $y_\tau = (\Delta c_\tau, \Delta d_\tau)$. 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

$$E\left[z_t|y^t\right] = \int z_t p\left(\theta, z_t|y^t\right) d\theta dz_t$$

Panel A in Figure C.13 shows the real-time estimates of the persistence parameter $E\left[\nu|y^t\right]$, 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_μ. As such, the growth rate of consumption is perceived as an i.i.d process, conditional on the Markov regime-switching state λ_{t+1}. After the great financial crisis the posterior mean of ν slightly increases with an end-of-sample estimate equal to $E\left[\nu|y^T\right] = 0.2$. 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 $E\left[E_\mu|y^t\right]$ is highly persistent and low volatile. However, this high persistence does not imply predictability in consumption growth since E_μ 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.[23] Top panel in Figure C.3 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.[24] 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 fil-

[23]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.

[24]They may be thought as standard maximum likelihood estimates, provided learning is unbiased, which is the case here.

tering. This may be interpreted as a rational expectations benchmark with no parameter uncertainty.

1.4.2 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 indefinitively 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 A.6 shows both the historical and the model-implied dynamics of the variance risk premium. As far as the unconditional moments are concerned, Tables A.5-C.8 provide the empirical moments and the corresponding statistics for the estimated model. 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.

Variance Risk Premium

Figure A.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 (A.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 C.3. 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 A.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.

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 A.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 A.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 A.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 γ from 5 to 2 slightly decreases the first component in (1.13), 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 C.8 provides the statistics results of

$$\frac{1}{k}\sum_{i=1}^{k}(r_{m,t+k}-r_{f,t+k})=\alpha_k+\beta^k_{VRP}VRP_{t,t+1}+\beta_{lpd}lpd_t+u_{t,t+k} \qquad k=1,...,24$$

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 heteroschedasticity and autocorrelation in the residuals.[25]

Top panel of Table C.8 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 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 pre-

[25] Here 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, a further Bayesian analysis would be redundant, at least for the explanatory purpose of this section.

dictive 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).[26]

1.4.3 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 λ_t. Indeed, the definition in (1.16), 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.

On top of its effect on the dispersion of the agent's belief about the economic outlook,

[26]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.

a jump in the uncertainty regime affects the weight put by the investor on current information. The so-called Kalman-Gain, indeed, is affected by the uncertainty regime and typically increases in bad states when the investor becomes more concerned about the state of the economy. Figure (C.4) shows the effect of an increasing perceived probability of a high uncertainty state.

An increasing perceived probability of being in a bad state increases the investor concerns about the current economic conditions, as well as the uncertainty about the economic outlook. This compounding effect reflects in an increasing variance risk premium under the model. The positive link between the uncertainty regime and the predictive variance of the drif is confirmed by the data, as shown in figure C.2.

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 C.2 it would be sensible to expect a positive relation between macroeconomic uncertainty and the index options premium.

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 $p(\lambda_{t+1} = H|y^{t+1}) - p(\lambda_{t+1} = H|y^{t})$.[27] 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 $Var[\Delta c_{t+1}|y^{t+1}] - Var[\Delta c_{t+1}|y^{t}]$.[28] 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.[29] 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 A.7 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.[30]

[27] 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 Π, (see Hamilton 1994).

[28] 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.

[29] 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.

[30] 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

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 C.8 shows the first component in (1.12), computed from the model with 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 C.10 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.

consumption expenditures.

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 C.9, the impact of the coefficient of risk aversion on the first component in (1.12) is relatively low. Therefore, the impact of relative risk aversion on the market variance risk premium is apparently flawed, consistent with the results statistics.

1.5 Conclusion

This chapter 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.

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 chapter 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.

Appendices

Appendix A

A.1 Data

A.1.1 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

$$\Delta c_{t+1} = \ln\left[\frac{C_{nd,t+1} + C_{s,t+1}}{C_{nd,t} + C_{s,t}}\right] \tag{A.1}$$

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.1.2 Stock market data

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

$$P_t = P_{t-1}\left(1 + VWRETX_t\right) \tag{A.2}$$

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.[1] The aggregate level of cash dividends is defined as

$$D_t = P_t^* \left[\frac{(1 + VWRETD_t)}{(1 + VWRETX_t)} - 1\right] \tag{A.3}$$

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 (PCEPI). Gross market index return and the realized equity premium are computed as

$$R_{t,t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} \qquad r_{t,t+1} - r_t^f = \ln\left(1 + R_{t,t+1}\right) - r_t^f$$

with r_t^f the ex-ante real risk free rate. Data are from CRSP and the sample period is 1990:01 - 2013:01, monthly.

A.1.3 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.

[1] 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.

A.1.4 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 $r_{f,t}$,

$$r_{f,t,t+1} = \hat{\alpha}_0 + \hat{\alpha}_1 r_{f,t} + \hat{\alpha}_2 \pi_{t-12,t} \tag{A.4}$$

A.1.5 The Implied and Realized Volatility Measures

The data on the VIX are from the Chicago Board of Options Exchange, i.e. CBOE.[2] 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 $E_t^{\mathbb{P}}[RV_{t,t+1}] = \alpha + \beta IV_{t-1,t} + \gamma RV_{t-1,t}$, 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 $E_t^{\mathbb{P}}[RV_{t,t+1}] = RV_{t-1,t}$ and $E_t^{\mathbb{P}}[RV_{t,t+1}] = \Psi(L)\epsilon_t$ with $\Psi(L)$ a lag-polynomial of order twelve (Bollerslev et al. 2009).

A.2 Variance Risk Premium

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

$$VRP_{t,t+\tau} = E_t^{\mathbb{Q}}[RV_{t,t+\tau}] - E_t^{\mathbb{P}}[RV_{t,t+\tau}] \tag{A.5}$$

[2] http://www.cboe.com/micro/vix/historical.aspx

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

$$d\mathbb{Q} = \frac{M_{t,t+\tau}}{E_t^{\mathbb{P}}\left[M_{t,t+\tau}\right]} d\mathbb{P} \qquad \text{such that} \qquad \frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{M_{t,t+\tau}}{E_t^{\mathbb{P}}\left[M_{t,t+\tau}\right]} \tag{A.6}$$

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

$$\begin{aligned} VRP_{t,t+\tau} &= E_t^{\mathbb{P}}\left[\frac{M_{t,t+\tau}}{E_t^{\mathbb{P}}\left[M_{t,t+\tau}\right]} RV_{t,t+\tau}\right] - E_t^{\mathbb{P}}\left[RV_{t,t+\tau}\right] \\ &= \frac{E_t^{\mathbb{P}}\left[M_{t,t+\tau} RV_{t,t+\tau}\right]}{E_t^{\mathbb{P}}\left[M_{t,t+\tau}\right]} - E_t^{\mathbb{P}}\left[RV_{t,t+\tau}\right] \end{aligned} \tag{A.7}$$

By using the law of iterated expectations

$$\begin{aligned} E\left[M_{t,t+\tau} RV_{t,t+\tau} | y^t\right] &= E\left[E\left[M_{t,t+\tau} RV_{t,t+\tau} | z_{t+1}, y^t\right] | y^t\right] \\ &= \sum_{i=1}^{k} E\left[M_{t,t+\tau} RV_{t,t+\tau} | \mu_{t+1}, \lambda_{t+1} = i, y^t\right] p\left(\lambda_{t+1} = i | y^t\right) \\ &= \sum_{i=1}^{k} E_{t,\mu}^{i}\left[M_{t,t+\tau} RV_{t,t+\tau}\right] p\left(\lambda_{t+1} = i | y^t\right) \end{aligned}$$

such as

$$\frac{E_t\left[M_{t,t+\tau} RV_{t,t+\tau}\right]}{E_t\left[M_{t,t+\tau}\right]} = \frac{\pi_{L,t+\tau} E_{t,\mu}^{L}\left[M_{t,t+\tau} RV_{t,t+\tau}\right] + \pi_{H,t+\tau}^{H} E_{t,\mu}^{H}\left[M_{t,t+\tau} RV_{t,t+\tau}\right]}{\pi_{L,t+\tau} E_{t,\mu}^{L}\left[M_{t,t+1}\right] + \pi_{H,t+\tau} E_{t,\mu}^{H}\left[M_{t,t+\tau}\right]}$$

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

$$E_{t,\mu}^{\mathbb{P}}\left[RV_{t,t+\tau}\right] = \pi_{L,t+\tau} E_{t,\mu}^{L}\left[RV_{t,t+\tau}\right] + \pi_{H,t+\tau} E_{t,\mu}^{H}\left[RV_{t,t+\tau}\right]$$

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

$$E_t\left[RV_{t,t+\tau}, M_{t,t+\tau}\right] = E_t\left[RV_{t,t+\tau}\right] E_t\left[M_{t,t+\tau}\right] + Cov_t\left[RV_{t,t+\tau}, M_{t,t+\tau}\right]$$

such that

$$\begin{aligned}
&\frac{E_t\left[RV_{t,t+\tau}, M_{t,t+\tau}\right]}{E_t\left[M_{t,t+\tau}\right]} = ...\\
&= \frac{\pi_{L,t+\tau}E^L_{t,\mu}\left[RV_{t,t+\tau}\right]E^L_{t,\mu}\left[M_{t,t+\tau}\right] + \pi_{H,t+\tau}E^H_{t,\mu}\left[RV_{t,t+\tau}\right]E^H_{t,\mu}\left[M_{t,t+\tau}\right]}{\pi_{L,t+\tau}E^L_{t,\mu}\left[M_{t,t+\tau}\right] + \pi_{H,t+\tau}E^H_{t,\mu}\left[M_{t,t+\tau}\right]} + \\
&+ \frac{\pi_{L,t+\tau}Cov^L_{t,\mu}\left[M_{t,t+\tau}, RV_{t,t+\tau}\right] + \pi_{H,t+\tau}Cov^H_{t,\mu}\left[M_{t,t+\tau}, RV_{t,t+\tau}\right]}{\pi_{L,t+\tau}E^L_{t,\mu}\left[M_{t,t+\tau}\right] + \pi_{H,t+\tau}E^H_{t,\mu}\left[M_{t,t+\tau}\right]}
\end{aligned} \tag{A.8}$$

Defining the preferences-adjusted probability of being in a high uncertainty state as

$$\tilde{\pi}_{H,t+\tau} = \left(\frac{E^H_{t,\mu}\left[M_{t,t+\tau}\right]}{\pi_{L,t+\tau}E^L_{t\mu}\left[M_{t,t+\tau}\right] + \pi_{H,t+\tau}E^H_{t,\mu}\left[M_{t,t+\tau}\right]} - 1\right)\pi_{H,t+\tau} \tag{A.9}$$

and collecting common terms we have

$$\begin{aligned}
VRP_{t,t+\tau} = {} & \tilde{\pi}_{H,t+\tau} \times \left(E^H_{t,\mu}\left[RV_{t,t+\tau}\right] - E^L_{t,\mu}\left[RV_{t,t+\tau}\right]\right) + ...\\
& + \frac{\left(\tau - \tilde{\pi}_{H,t+\tau}\right)}{E^L_{t,\mu}\left[M_{t,t+\tau}\right]}Cov^L_{t,\mu}\left[RV_{t,t+\tau}, M_{t,t+\tau}\right] + \frac{\tilde{\pi}_{H,t+\tau}}{E^H_{t,\mu}\left[M_{t,t+\tau}\right]}Cov^H_{t,\mu}\left[RV_{t,t+\tau}, M_{t,t+\tau}\right]
\end{aligned}$$

now defining the last covariance terms κ_t the definition in the main text follows.

A.3 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 $\theta = (E_\mu, \nu, \mu_d, \phi, \sigma^2_c, \sigma^2_d, p_{LL}, p_{HH}, \sigma^2_L, \sigma^2_H)$. The investor holds initial beliefs over the states and parameters $p(\theta, z_t|y^t) = p(z_t|\theta, y^t)\, p(\theta|y^t)$ and updates them via sequential Bayes' rule, in which $y_\tau = (\Delta c_\tau, \Delta d_\tau)$ and $y^t = (y_1, ..., y_t)$.

The learning scheme occurs in two steps: A prediction step

$$p(z_{t+1}, \theta|y^t) = \int p(z_{t+1}|z_t, \theta)p(\theta, z_t|y^t)dz_t \tag{A.10}$$

then belief updating via the predictive likelihood $p\left(y_{t+1}|z_{t+1}, \theta\right)$

$$p(z_{t+1}, \theta|y^{t+1}) \propto p(y_{t+1}|z_{t+1}, \theta)p(z_{t+1}, \theta|y^t) \tag{A.11}$$

which shows the recursive nature of Bayesian updating, as $p(z_{t+1}, \theta|y^{t+1})$ is functionally dependent on $p(z_t, \theta|y^t)$. The main issue is characterize $p(z_t, \theta|y^t)$ 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 κ_t and the states κ_t^x such that

$$p\left(\theta, \mu_{t+1}, \lambda_{t+1}, \kappa_{t+1}, \kappa_{t+1}^x|y^{t+1}\right) = p\left(\theta|\kappa_{t+1}\right) p\left(\mu_{t+1}|\kappa_{t+1}^x, \lambda_{t+1}, y^{t+1}\right) p\left(\kappa_{t+1}, \kappa_{t+1}^x, \lambda_{t+1}|y^{t+1}\right)$$

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

$$\lambda_{t+1} \sim p(\lambda_{t+1}|\kappa_t^x, \lambda_t, \theta, \Delta c_{t+1}) \propto p(\Delta c_{t+1}|\mu_t, \theta, \lambda_{t+1})p(\lambda_{t+1}|\lambda_t, \theta)$$

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

$$\begin{aligned} f_{t+1} &= (1-\nu)\, E_\mu + \nu m_t \\ Q_{t+1} &= R_{t+1} + \sigma_c^2 \end{aligned}$$

with $R_{t+1} = \nu^2 C_t + \sigma^2_{\mu,\lambda_{t+1}}$ the predictive variance of μ_{t+1} (see West and Harrison 1997). Given the current information y_{t+1} and the simulated uncertainty state λ_{t+1} the drift μ_{t+1} is propagated from the conditional sufficient statistics κ_{t+1} and κ^x_{t+1}

$$\kappa_{t+1} = \mathcal{K}\left(\kappa_t, z_{t+1}, y_{t+1}\right)$$
$$\kappa^x_{t+1} = \mathcal{K}^x\left(\kappa^x_t, \theta, \lambda_{t+1}, y_{t+1}\right)$$

in which $\mathcal{K}^x(\cdot)$ evolves as $\kappa^x_{t+1} = (m_{t+1}, C_{t+1})$

$$m_{t+1} = a_{t+1} + A_{t+1} e_{t+1}$$
$$C_{t+1} = R_{t+1} - A_{t+1} Q^{-1}_{t+1} A_{t+1}$$

with $a_{t+1} = (1-\nu) E_\mu + \nu m_t$ the predictive mean, $e_{t+1} = \Delta c_{t+1} - E_t \Delta c_{t+1}$ the forecasting error and $A_{t+1} = R_{t+1} Q^{-1}_{t+1}$ 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 $\kappa_{t+1} = \mathcal{K}\left(\kappa_t, z_{t+1}, y_{t+1}\right)$ is updated. The agent's prior about the drift parameters is a standard normal-inverse gamma prior

$$p\left(\beta_\mu | \sigma^2_{\mu,\lambda_t}, \kappa_t\right) \sim N\left(b_t, \sigma^2_{\lambda_t} B^{-1}_t\right)$$
$$p\left(\sigma^2_{\lambda_{\mu,t}=i} | \kappa_t\right) \sim IG\left(v_{i,t}/2, V_{i,t}/2\right) \qquad i = H, L$$

in which b_t, B_t represent respectively a location and scale parameter, $v_{i,t}$ counts the degrees of freedom for the ith 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

$$p\left(\beta_\mu | \sigma^2_{\lambda_{t+1}}, \kappa_{t+1}\right) \sim N\left(b_{t+1}, \sigma^2_{\mu,\lambda_{t+1}} B^{-1}_{t+1}\right)$$
$$p\left(\sigma^2_{\lambda_{t+1}=i} | \kappa_{t+1}\right) \sim IG\left(v_{i,t+1}/2, V_{i,t+1}/2\right)$$

in which the specific Kalman-like recursion evolves as

$$\begin{aligned}
B_{t+1}^{-1} &= B_t^{-1} + Z_{t+1} Z_{t+1}' \\
B_{t+1}^{-1} b_{t+1} &= B_t^{-1} b_t + Z_t \mu_{t+1} \\
v_{i,t+1} &= v_{i,t} + 1 \qquad \text{for} \qquad i = H, L \\
V_{i,t+1} &= V_{i,t} + \left[\left(\mu_{t+1} - Z_t^\top b_{t+1} \right) \mu_{t+1} + \left(b_t - b_{t+1} \right)^\top B_t^{-1} b_t \right] \mathbb{I}_{\lambda_{t+1} = i}
\end{aligned}$$

with $Z_t = [1, \mu_t]$, and $\beta_\mu = [(1 - \nu) E_\mu, \nu]$. The prior beliefs at time t of the idiosyncratic risk σ_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 Δc_{t+1} and given μ_{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

$$\begin{aligned}
n_{t+1} &= n_t + 1 \\
N_{t+1} &= N_t + \left(\Delta c_{t+1} - \mu_{t+1} \right)^2
\end{aligned}$$

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

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

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 μ_{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

$$\begin{aligned} G_{t+1}^{-1} &= G_t^{-1} + X_{t+1}X_{t+1}' \\ G_{t+1}^{-1}g_{t+1} &= G_t^{-1}g_t + X_t\Delta d_{t+1} \\ d_{t+1} &= d_t + 1 \\ D_{t+1} &= D_t + \left(\Delta d_{t+1} - X_t^\top g_{t+1}\right)\Delta d_{t+1} + \left(g_t - g_{t+1}\right)^\top G_t^{-1} g_t \end{aligned}$$

with $X_t = [1, \mu_{t+1} - E_\mu]'$. Finally the agent's beliefs about the transition matrix Π are updated as follows

$$p_{HH} \sim \frac{\Gamma\left(p_{1,t} + p_{2,t}\right)}{\Gamma\left(p_{1,t}\right)\Gamma\left(p_{2,t}\right)} p^{p_{1,t}-1}\left(1-p\right)^{p_{2,t}-1} \mathbb{I}_{]0,1[}(p)$$
$$p_{LL} \sim \frac{\Gamma\left(q_{1,t} + q_{2,t}\right)}{\Gamma\left(q_{1,t}\right)\Gamma\left(q_{2,t}\right)} q^{q_{1,t}-1}\left(1-q\right)^{q_{2,t}-1} \mathbb{I}_{]0,1[}(q)$$

with

$$p_{1,t+1} = p_{1,t} + \mathbb{I}_{(\lambda_{t+1}=H, \lambda_t=H)} \qquad q_{1,t+1} = q_{1,t} + \mathbb{I}_{(\lambda_{t+1}=L, \lambda_t=L)}$$
$$p_{2,t+1} = p_{2,t} + \mathbb{I}_{(\lambda_{t+1}=H, \lambda_t=L)} \qquad q_{2,t+1} = q_{2,t} + \mathbb{I}_{(\lambda_{t+1}=L, \lambda_t=H)}$$

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

$$E\left[p_{HH}|y^t,\kappa_t\right] = \frac{p_{1,t+1}}{p_{1,t+1}+p_{2,t+1}} \tag{A.12}$$

which is an increasing function of $p_{1,t+1}$.

A.3.1 Hypothesis Testing

Model assessment is done by comparing the marginal posterior models' probabilities, which are computed integrating out both states and parameter uncertainty.[3] Given the specific prior model probability $p(\mathcal{M}_i)$ the posterior probability of model i is computed as

$$p\left(\mathcal{M}_i|y^t\right) = \frac{p\left(y^t|\mathcal{M}_i\right)p(\mathcal{M}_i)}{\sum_{i=1}^{I} p\left(y^t|\mathcal{M}_i\right)p(\mathcal{M}_i)} \tag{A.13}$$

The real-time nature of the learning scheme allows to compute the marginal likelihood recursively as

$$p\left(y^t|\mathcal{M}_i\right) = p\left(y_t|y^{t-1},\mathcal{M}_i\right)p\left(y^{t-1}|\mathcal{M}_i\right) \tag{A.14}$$

where

$$p\left(y_t|y^{t-1},\mathcal{M}_i\right) = \int p\left(y_t|y^{t-1},\theta,z_{t-1},\mathcal{M}_i\right)p\left(z_{t-1},\theta|y^{t-1},\mathcal{M}_i\right)dz_{t-1}d\theta \tag{A.15}$$

The recursive predictive probability $p\left(y_t|y^{t-1},\mathcal{M}_i\right)$ is computed from the M particles as

$$p\left(y_t|y^{t-1},\mathcal{M}_i\right) \approx \frac{1}{M}\sum_{m=1}^{M} p\left(y_t\middle|\left(\theta,\kappa_{t-1},\kappa_{t-1},^{x},z_{t-1}\right)^{(m)},y^{t-1},\mathcal{M}_i\right) \tag{A.16}$$

[3]Notice that, by integrating out parameter uncertainty the marginal likelihood punishes needlessly complicated models.

The null hypothesis $\mathcal{H}_0 : \nu = 0$ against the alternative $\mathcal{H}_1 : \nu \neq 0$ is tested by using standard Bayes factors. Since the model with $\nu = 0$ is nested in the more general unrestricted version the Bayes factors $\mathcal{BF}^t_{0,1}$ might be approximated via the Savage-Dickey density ratio

$$\mathcal{BF}^t_{0,1} = \frac{p\left(\mathcal{H}_0|y^t\right)}{p\left(\mathcal{H}_1|y^t\right)} = \frac{p\left(\nu = 0|y^t, \mathcal{M}_1\right)}{p\left(\nu = 0|\mathcal{M}_1\right)}$$

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

$$p\left(\nu = 0|y^t, \mathcal{M}_1\right) = \int p\left(\nu = 0|\theta_{[-\nu]}, \kappa_t, \kappa^x_t\right) p\left(\theta_{[-\nu]}, \kappa_t, \kappa^x_t|y^t\right) d\left(\theta_{[-\nu]}, \kappa_t, \kappa^x_t\right) \qquad \text{(A.17)}$$

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

$$p\left(\nu = 0|y^t, \mathcal{M}_1\right)^N = \frac{1}{M}\sum_{m=1}^{M} p\left(\nu = 0|\left(\theta_{[-\nu]}, \kappa_t, \kappa^x_t\right)^{(m)}\right) \qquad \text{(A.18)}$$

Assuming a priori that $p\left(\mathcal{H}_0\right) = p\left(\mathcal{H}_1\right)$, the posterior probability of the null hypothesis can be computed from the Bayes factor $\mathcal{BF}^t_{0,1}$ as

$$p\left[\mathcal{H}_0|y^t\right] = \frac{\mathcal{BF}^t_{0,1}}{1 + \mathcal{BF}^t_{0,1}} \qquad \text{(A.19)}$$

A.4 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.[4] The location parameter for $p\left(E_\mu\right)$ is set

[4]http://www.econ.yale.edu/shiller/data.htm

to be 0.17% on a monthly basis. Likewise the location of $p(\mu_d)$ is set to be equal to 0.09% on a monthly basis.[5] The location of the persistent parameter is set to be $E\left[\nu|y^0\right] = 0.97$, consistent with the long-run risk literature (Bansal and Yaron 2004, Bansal, Kiku, and Yaron 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 $E\left[\phi\right] = 3.5$ which is in line with Bansal and Yaron (2004), Lettau et al. (2008) and Abel (1999) among the others. Finally the prior mean of σ_c^2 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 $\sigma_{\mu,t+1}^2$ is assumed to be higher a priori under $\lambda_{t+1} = H$ for identification purposes. Table A.8 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

$$p\left(y^T\right) = \prod_{t=1}^{T} p\left(y_t|y_{t-1}\right) p\left(y_0\right) = \frac{1}{N^T} \prod_{t=1}^{T} \sum_{m=1}^{M} p\left(y_t|\left(z_t,\theta\right)^{(i)}\right) \tag{A.20}$$

in which $\left(z_t,\theta\right)^{(i)}$ are particles from the predictive $p\left(z_t,\theta|y^{t-1}\right)$. Table A.9 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 A.12 shows the

[5] 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.

results. I set the prior model probability $p(\mathcal{M}_i)$ of each model to be 1/2.

The blue area shows the marginal posterior probability of the model with two regimes. The single regime model is quickly rejected by the agent that updates belief based on current real per capita consumption.[6] 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 A.13 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.

A.4.1 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 A.10 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. $E\left[E_\mu|y^T\right] = 0.2110$.

The long-run growth rate of the aggregate dividend, i.e. $E\left[\mu_d|y^T\right] = 0.226$ 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 $E\left[P_{HH}|y^T\right] = 0.512$ as opposed to $E\left[P_{LL}|y^T\right] = 0.8688$. 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 $E\left[\pi_H|y^T\right] = 0.22$ and

[6] 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.

$E\left[\pi_L|y^T\right] = 0.78$.[7] The compounding effect of $\sigma^2_{\mu,\lambda_{t+1}=H}$ and $E\left[p_{HH}|y^T\right] = 0.51$ indicates that macroeconomic uncertainty shocks are infrequent but not rare, and large but do consistent with the rare disaster literature.

A.5 Numerical Solution of the Model

The dynamics of economic fundamentals is given by

$$\begin{aligned}
\Delta c_{t+1} &= \mu_{t+1} + \sigma_c \epsilon_{c,t+1} & \epsilon_{c,t+1} &\sim N(0,1) \\
\mu_{t+1} &= (1-\nu)E_\mu + \nu\mu_t + \sigma_{\mu,\lambda_{t+1}}\epsilon_{\mu,t+1} & \epsilon_{\mu,t+1} &\sim N(0,1) \\
\Delta d_{t+1} &= \mu_d + \phi\left(\mu_{t+1} - E_\mu\right) + \sigma_d \epsilon_{d,t+1} & \epsilon_{d,t+1} &\sim N(0,1)
\end{aligned}$$

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

$$\Pi' = \begin{pmatrix} p_{LL} & 1-p_{HH} \\ 1-p_{LL} & p_{HH} \end{pmatrix} \tag{A.21}$$

The regime changes are assumed to be indipendent 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

[7] 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

$$Dur_i = \frac{1}{1-p_{ii}}$$

the standard form

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

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 ket implication is that, ex-ante, beliefs on states and parameters are seen as if they will remain constant indefinitively 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 $\kappa_{t+1}^x = (m_{t+1}, C_{t+1})$ 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 $\Delta c_{t+1} = g(\kappa_t^x, \lambda_{t+1})$. Further, it is necessary to bound the support of the conditional beliefs on the expected growth rate of consumption; $m_t \in (\underline{m}, \overline{m})$ 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

$$\begin{aligned}
P^C\left(\kappa_t^x, \lambda_t\right)^\theta &= E\left[\beta^\theta \exp^{(1-\gamma)\Delta c_{t+1}(\kappa_t^x,\lambda_{t+1})}\left(1+P^C(\kappa_{t+1}^x,\lambda_{t+1})\right)^\theta \middle| \kappa_t^x, \lambda_t\right] \\
&= \beta^\theta E\left[E\left[\exp^{(1-\gamma)\Delta c_{t+1}(\kappa_{t+1}^x)}\left(1+P^C(\kappa_{t+1}^x,\lambda_{t+1})\right)^\theta \middle| \kappa_{t+1}^x, \lambda_{t+1}\right] \middle| \kappa_t^x, \lambda_t\right] \\
&= \beta^\theta E\left[\exp^{(1-\gamma)m_{t+1}+\frac{1}{2}(1-\gamma)^2 C_{t+1}}\left(1+P^C(\kappa_{t+1}^x,\lambda_{t+1})\right)^\theta \middle| \kappa_t^x, \lambda_t\right] \\
&= \beta^\theta \sum_{\lambda_{t+1}=L}^{H} p(\lambda_{t+1}|\lambda_t,\kappa_t^x)\exp^{(1-\gamma)m_{t+1}+\frac{1}{2}(1-\gamma)^2 C_{t+1}} \times ... \\
&... \left(1+P^C(\kappa_{t+1}^x,\lambda_{t+1})\right)^\theta
\end{aligned}$$

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 λ_t and the exogenous shocks in consumption $\epsilon_{c,t+1}$ are conditionally indipendent. Notice that the predictive probability of the ith regime, as well as the conditional moments m_{t+1}, C_{t+1} are found integrating out parameter uncertainty,

$$p(\lambda_{t+1}|\lambda_t,\kappa_t^x) = \int p(\lambda_{t+1}|\lambda_t,\theta,\kappa_t^x)p(\lambda_t,\theta|\kappa_t^x)d\theta$$

where θ 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

$$E\left[\Delta d_{t+1}|y^{t+1}\right] = \mu_d + E\left[\mu_{t+1} - E_\mu|y^{t+1}\right] = \mu_d + \nu\left(m_{t+1} - E_\mu\right)$$
$$Var\left[\Delta d_{t+1}|y^{t+1}\right] = Var\left[\mu_{t+1} - E_\mu|y^{t+1}\right] + \sigma_d^2 = C_{t+1} + \sigma_d^2$$

As such the price-dividend ratio can be solved as the fixed point of

$$\begin{aligned}
P^D\left(\kappa_t^x, \lambda_t\right) &= E\left[\beta^\theta \exp^{(-\gamma\Delta c_{t+1}(\lambda_{t+1})+\Delta d_{t+1}(\lambda_{t+1}))}\left(\frac{1+P^C(\kappa_{t+1}^x, \lambda_{t+1})}{P^C(\kappa_t^x, \lambda_t)}\right)^{\theta-1}\times\right. \\
&\left.\times\ P^D\left(\lambda_{t+1}, \kappa_{t+1}^x + 1\right)\middle|\kappa_t^x, \lambda_t\right] \\
&= E\left[E\left[\beta^\theta \exp^{(-\gamma\Delta c_{t+1}(\lambda_{t+1})+\Delta d_{t+1}(\lambda_{t+1}))}\left(\frac{1+P^C(\kappa_{t+1}^x, \lambda_{t+1})}{P^C(\kappa_t^x, \lambda_t)}\right)^{\theta-1}\times ...\right.\right. \\
&\left. ...\ P^D\left(\lambda_{t+1}, \kappa_{t+1}^x\right)|\kappa_t^x, \lambda_{t+1}, \lambda_t|\kappa_t^x, \lambda_t\right] \\
&= \beta^\theta \sum_{\lambda_{t+1}=L}^{H} p\left(\lambda_{t+1}|\lambda_t, \kappa_t^x\right)\exp^{(-\gamma\Delta c_{t+1}(\lambda_{t+1})+\Delta d_{t+1}(\lambda_{t+1}))} \\
&\times\left(\frac{1+P^C(\kappa_{t+1}^x, \lambda_{t+1})}{P^C(\kappa_t^x, \lambda_t)}\right)^{\theta-1}\times P^D\left(\lambda_{t+1}, \kappa_{t+1}^x + 1\right)
\end{aligned}$$

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.[8]

A.6 Tables and Figures

[8]The number of discretizing points G is chosen to be odd such that m_t is in the middle point of the grid.

Table A.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. $IV_{t,t+1} = VIX^2_{t,t+1}/12$ 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.

	Sub Sample 1990:01-2007:12					**Full Sample 1990:01-2013:01**				
	Mkt	$\mathbf{IV_{t,t+1}}$	$\mathbf{RV_{t,t+1}}$	$\mathbf{E_tRV_{t,t+1}}$	$\mathbf{VRP_{t,t+1}}$	**Mkt**	$\mathbf{IV_{t,t+1}}$	$\mathbf{RV_{t,t+1}}$	$\mathbf{E_tRV_{t,t+1}}$	$\mathbf{VRP_{t,t+1}}$
Mean	0.51	33.48	15.41	16.98	15.82	0.42	39.79	21.32	21.39	18.36
Median	1.02	26.74	9.85	12.57	10.76	1.02	31.28	12.41	15.50	12.69
Std. Dev	4.06	23.65	15.80	13.11	17.40	4.54	35.61	37.47	24.44	20.54
Skewness	-0.82	1.98	2.58	1.84	2.37	-0.89	3.32	7.82	8.15	2.16
Kurtosis	4.56	9.00	11.64	7.05	11.89	4.91	19.00	86.70	92.92	9.32
AR(1)	0.04	0.78	0.68	0.80	0.56	0.11	0.80	0.65	0.63	0.57
AR(2)	-0.01	0.60	0.54	0.63	0.39	-0.02	0.59	0.40	0.37	0.42
min	-17.60	9.04	1.86	3.88	-2.53	-20.49	9.05	1.87	9.15	-26.91
max	9.76	163.39	111.46	73.01	125.30	10.79	298.90	479.58	325.81	125.30
Jarque-Bera	47.34	470.49	921.08	272.61	923.81	82.011	3388.4	81626	95343.6	663.75

Table A.2: Variance Risk Premium and Long-Run Predictive Regressions

This table reports a set of predictive regressions with $k = 1, ..., 24$ months as forecasting horizon. The indipendent 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 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. 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.

Predictive Regressions: Standard Bayesian								
Horizon (Months)	**1**	**3**	**6**	**9**	**12**	**15**	**18**	**24**
Intercept	−0.207	-0.203	-0.004	0.152	0.197	0.233	0.267	0.291
	[-0.74 0.38]	[-0.54 -0.15]	[-0.29 0.26]	[-0.05 0.37]	[0.01 0.39]	[0.05 0.40]	[0.09 0.42]	[0.11 0.43]
$\mathbf{VRP_{t,t+1}}$	0.347	0.489	0.249	0.167	0.144	0.119	0.097	0.084
	[0.12 0.55]	[0.32 0.58]	[0.15 0.35]	[0.09 0.25]	[0.07 0.22]	[0.05 0.18]	[0.04 0.16]	[-0.01 0.13]
Adj $\mathbf{R^2}$(50%)	2.36	6.82	5.21	3.41	3.20	2.77	2.30	0.82
SR Max/SR Unc	1.94	3.05	2.69	2.24	2.18	2.07	1.92	1.38

Predictive Regressions: Robust Bayesian								
Horizon (Months)	**1**	**3**	**6**	**9**	**12**	**15**	**18**	**24**
Intercept	-0.333	-0.374	-0.131	0.084	0.116	0.159	0.189	0.211
	[-1.11 0.45]	[-0.86 0.11]	[-0.48 0.23]	[-0.22 0.33]	[-0.14 0.36]	[-0.06 0.37]	[-0.03 0.39]	[0.02 0.38]
$\mathbf{VRP_{t,t+1}}$	0.32	0.361	0.252	0.164	0.141	0.126	0.109	0.073
	[0.02 0.62]	[0.18 0.54]	[0.13 0.38]	[0.06 0.27]	[0.03 0.24]	[0.02 0.20]	[-0.01 0.18]	[-0.02 0.16]
Adj $\mathbf{R^2}$(50%)	2.21	5.24	4.56	2.94	2.52	2.20	1.84	0.31
SR Max/SR Unc	1.89	2.74	2.55	2.10	1.98	1.89	1.76	1.16

Table A.3: Variance Risk Premium and Standard Predictors

This table reports a set of one-step-ahead forecasting regressions. The indipendent 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).

Dependent $\mathbf{r_{m,t+1} - r_{f,t+1}}$	Models													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Intercept	-0.69	-0.22	-0.34	-0.31	-0.29	-0.02	-0.14	0.14	-0.13	4.90	4.08	11.6**	1.39	13.34***
	(-1.52)	(-0.61)	(-0.73)	(-0.57)	(-0.53)	(-0.06)	(-0.51)	(0.38)	(-0.38)	(1.07)	(1.49)	(2.18)	(1.44)	(2.67)
$\mathbf{VRP^{(\mathcal{M}_0)}_{t,t+1}}$	0.52***			0.57***	0.45***									
	(3.02)			(3.07)	(2.35)									
$\mathbf{VRP^{(\mathcal{M}_1)}_{t,t+1}}$		0.36***				0.38***	0.48***							
		(3.49)				(3.87)	(3.21)							
$\mathbf{VRP^{(\mathcal{M}_2)}_{t,t+1}}$			0.32***					0.49***	0.51***	0.33*	0.32**		0.56***	0.60***
			(2.97)					(2.29)	(2.74)	(1.95)	(2.15)		(2.92)	(3.21)
$\mathbf{IV}_{t,t+1}$				-0.12		0.11		-0.19						
				(-1.01)		(0.67)		(-1.27)						
$\mathbf{RV}_{t-1,t}$					-0.07		-0.21		−0.28***					
					(-1.04)		(-1.60)		(-2.39)					
$\mathbf{log(P_t/D_t)}$										-1.77				
										(-1.72)				
$\mathbf{log(P_t/E_t)}$											−1.67*	−2.80*		−3.40**
											(-1.81)	(-1.93)		(-2.42)
$\mathbf{TERM_t}$												-0.21	0.22	-0.14
												(-0.80)	(0.76)	(-0.48)
$\mathbf{DEF_t}$												−1.93*	−2.42**	−3.10***
												(-1.80)	(-2.40)	(-1.06)
$\mathbf{RealRf_t}$												-1.68	-1.19	-1.20
												(-0.56)	(-0.22)	(-0.23)
Adj $\mathbf{R^2}$(%)	3.6	3.2	2.2	3.7	3.9	3.4	3.0	2.5	5.4	4.7	6.2	3.8	4.9	7.9
SR Max/SR Unc	2.3	2.2	1.8	2.3	2.4	2.2	2.1	1.9	2.7	2.5	2.9	2.3	2.6	3.3

Table A.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.

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

Table A.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. 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

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

Panel B: Rational Expectations Benchmark

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

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

Moments	Data		$\gamma = 5$					$\gamma = 2$				
	Estimate	(St.Err)	Mean	(St.Err)	2.5%	50%	97.5%	Mean	(St.Err)	2.5%	50%	97.5%
VRP												
$E_T(VRP_{t+1,t})$	18.36	(2.44)	11.48	(1.62)	9.23	11.82	14.61	11.66	(1.90)	8.60	11.64	14.86
$\sigma_T(VRP_{t+1,t})$	20.54	(4.57)	11.20	(2.43)	7.11	11.30	15.06	10.22	(2.72)	5.54	10.27	14.59
skewness	2.16	(0.77)	1.70	(0.44)	1.11	1.65	2.51	1.67	(0.43)	1.07	1.61	2.46
kurtosis	9.32	(5.60)	4.35	(1.07)	2.74	4.75	6.01	4.02	(0.97)	2.40	4.48	6.53
$\rho_T(VRP_{t+1,t}, VRP_{t-1,t})$	0.57	(0.05)	0.38	(0.06)	0.27	0.39	0.47	0.30	(0.07)	0.17	0.30	0.40

Table A.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 $k = 1, ..., 12$. The indipendent variables are the current model-implied variance risk premium and log price-dividend ratio. Regression estimates are robust for autocorrelation and heteroschedasticity (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							
Months		**Data**				**Model**							
k	β^k_{VRP}	**(t-stat)**	β^k_{lpd}	**(t-stat)**	R^2_{adj}	$\hat{\beta}^k_{VRP}$	β^k_{VRP}(50%)	**t-stat VRP**	$\hat{\beta}^k_{lpd}$	β^k_{lpd}(50%)	**t-stat lpd**	$\hat{R}^2_{adj}$	R^2_{adj}(50%)
1	0.56	(5.37)	-1.77	(-1.72)	5.51	0.61	0.62	(2.26)	-1.84	-1.82	(-0.81)	2.21	1.99
3	0.45	(7.97)	-1.75	(-1.94)	10.7	0.32	0.33	(2.40)	-2.02	-1.99	(-1.23)	3.74	3.70
6	0.28	(5.11)	-2.71	(-2.11)	11.6	0.18	0.17	(1.96)	-2.35	-2.33	(-1.99)	5.32	5.72
9	0.17	(3.04)	-2.95	(-2.48)	13.0	0.12	0.13	(1.26)	-2.54	-2.53	(-2.41)	8.87	8.84
12	0.13	(2.51)	-3.15	(-2.87)	16.3	0.03	0.03	(0.11)	-3.06	-3.04	(-2.65)	10.5	10.2
						Panel B: Rational Expectations Benchmark							
Months		**Data**				**Model**							
k	β^k_{VRP}	**(t-stat)**	β^k_{lpd}	**(t-stat)**	R^2_{adj}	$\hat{\beta}^k_{VRP}$	β^k_{VRP}(50%)	**t-stat VRP**	$\hat{\beta}^k_{lpd}$	β^k_{lpd}(50%)	**t-stat lpd**	$\hat{R}^2_{adj}$	R^2_{adj}(50%)
1	0.56	(5.37)	-1.77	(-1.72)	5.51	0.20	0.18	(1.89)	-0.28	-0.32	(-0.03)	1.12	1.00
3	0.45	(7.97)	-1.75	(-1.94)	10.7	0.14	0.13	(1.52)	-1.01	-0.81	(-0.97)	1.54	1.32
6	0.28	(5.11)	-2.71	(-2.11)	11.6	0.13	0.12	(1.35)	-1.98	-1.94	(-1.23)	1.86	1.45
9	0.17	(3.04)	-2.95	(-2.48)	13.0	0.04	0.02	(0.31)	-2.21	-2.25	(-1.92)	2.14	1.95
12	0.13	(2.51)	-3.15	(-2.87)	16.3	0.02	0.01	(0.25)	-2.38	-2.36	(-2.14)	2.32	2.25
						Panel C: Rational Expectations Benchmark ($\nu = 0.98$)							
Months		**Data**				**Model**							
k	β^k_{VRP}	**(t-stat)**	β^k_{lpd}	**(t-stat)**	R^2_{adj}	$\hat{\beta}^k_{VRP}$	β^k_{VRP}(50%)	**t-stat VRP**	$\hat{\beta}^k_{lpd}$	β^k_{lpd}(50%)	**t-stat lpd**	$\hat{R}^2_{adj}$	R^2_{adj}(50%)
1	0.56	(5.37)	-1.77	(-1.72)	5.51	0.36	0.38	(1.69)	-0.02	-0.02	(-0.01)	1.23	1.03
3	0.45	(7.97)	-1.75	(-1.94)	10.7	0.26	0.28	(2.76)	-1.40	-1.46	(-1.00)	2.00	2.02
6	0.28	(5.11)	-2.71	(-2.11)	11.6	0.15	0.10	(1.94)	-2.16	-2.08	(-1.71)	3.84	3.82
9	0.17	(3.04)	-2.95	(-2.48)	13.0	0.05	0.04	(1.47)	-2.50	-2.99	(-2.00)	5.22	5.00
12	0.13	(2.51)	-3.15	(-2.87)	16.3	0.03	0.02	(0.97)	-3.07	-3.00	(-2.71)	7.12	7.17

Table A.7: 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.

	Models									
Indipendent: $\mathbf{VRP_{t,t+1} - VRP_{t-1,t}}$	1	2	3	4	5	6	7	8	9	10
$\mathbf{p}(\lambda_{\mathbf{t+1}}\vert\mathbf{y^{t+1}}) - \mathbf{p}(\lambda_{\mathbf{t+1}}\vert\mathbf{y^{t}})$	0.188**	0.145**	0.159**	0.126*	0.182*	0.132*	0.183*	0.166*	0.141*	0.094*
	(2.028)	(1.993)	(2.001)	(1.849)	(1.846)	(1.799)	(1.899)	(1.871)	(1.892)	(1.859)
Other Beliefs:										
$\mathbf{E}[\mathbf{\Delta c_{t+1}}\vert\mathbf{y^{t+1}}] - \mathbf{E}[\mathbf{\Delta c_{t+1}}\vert\mathbf{y^{t}}]$		−0.161*		−0.133*		−0.097*			−0.123**	−0.159*
		(-0.165)		(-1.801)		(-1.781)			(-0.201)	(-1.866)
$\mathbf{Var}[\mathbf{\Delta c_{t+1}}\vert\mathbf{y^{t+1}}] - \mathbf{Var}[\mathbf{\Delta c_{t+1}}\vert\mathbf{y^{t}}]$			0.255**	0.236**		0.224**			0.197***	0.189*
			(1.985)	(2.156)		(1.925)			(2.422)	(1.895)
Controls:										
$\mathbf{\Delta c_{t+1}}$									-0.108	-0.153
									(-1.359)	(-1.511)
$\mathbf{\Delta c_{t}}$									-0.021	-0.101
									(-0.348)	(-1.112)
Adj $\mathbf{R^2}$(%)	5.39	6.20	7.71	9.25	9.27	14.8	10.1	10.2	10.3	18.7

Table A.8: Prior Calibration

This table reports the hyper-parameters for the model priors 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.

Parameter	Location	Scale
ν	0.980	0.50
$\mathbf{E}_{\mu}$	0.170	0.50
$\mu_{\mathbf{d}}$	0.090	0.50
ϕ	3.000	0.50
$\mathbf{P_{LL}}$	0.950	0.50
$\mathbf{P_{HH}}$	0.950	0.50
$\sigma^{\mathbf{2}}_{\mu,\lambda_{\mathbf{t}}=\mathbf{H}}$	0.263	0.50
$\sigma^{\mathbf{2}}_{\mu,\lambda_{\mathbf{t}}=\mathbf{L}}$	0.052	0.50

Table A.9: 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.

	Real-Time Structural Learning		Rational Expectations Benchmark	
	One Regime	Two Regimes	One Regime	Two Regimes
Marginal Likelihood	-118.301	**-107.110**	-123.142	-117.213
Bayes Factor	> 100		> 100	> 100

Table A.10: Parameters Estimates

This table reports the paramater estimates of the dynamics

$$\begin{aligned}
\Delta c_{t+1} &= \mu_{t+1} + \sigma_c \epsilon_{c,t+1} \\
\Delta d_{t+1} &= \mu_d + \phi\left(\mu_{t+1} - E_\mu\right) + \sigma_d \epsilon_{d,t+1} \\
\mu_{t+1} &= (1-\nu)E_\mu + \nu \mu_t + \sigma_{\mu,\lambda_{t+1}} \epsilon_{\mu,t+1} \qquad [\epsilon_{c,t+1}, \epsilon_{\mu,t+1}, \epsilon_{d,t+1}]' \sim N(0, I_3)
\end{aligned}$$

where $\lambda_t = i$, for $i = H, L$ follows a transition probability

$$p\left(\lambda_{t+1} = H | \lambda_t = H, \theta\right) = p_{HH} \qquad \text{and} \qquad p\left(\lambda_{t+1} = L | \lambda_t = L, \theta\right) = p_{LL}$$

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.

Panel A: End-of-Sample Estimates					
Parameter	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**
ν	0.2081	(0.0571)	0.1206	0.2007	0.3022
$\mathbf{E}_\mu$	0.1910	(0.0212)	0.1710	0.2110	0.2320
$\sigma^{\mathbf{2}}_{\mu,\lambda_\mathbf{t}=\mathbf{L}}$	0.0321	(0.0036)	0.0310	0.0331	0.0353
$\sigma^{\mathbf{2}}_{\mu,\lambda_\mathbf{t}=\mathbf{H}}$	0.2416	(0.0732)	0.1507	0.2279	0.3820
$\mu_\mathbf{d}$	0.2265	(0.0967)	0.0807	0.2267	0.3816
ϕ	2.6212	(0.7762)	1.3492	2.6321	3.9012
$\mathbf{p_{HH}}$	0.5129	(0.0005)	0.5121	0.5129	0.5136
$\mathbf{p_{LL}}$	0.8688	(0.0002)	0.8685	0.8688	0.8690
$\pi_\mathbf{L}$	0.7747	(0.0003)	0.7743	0.7747	0.7751
$\pi_\mathbf{H}$	0.2253	(0.0004)	0.2257	0.2253	0.2249
$\mathbf{Dur_L}$	7.0577	(0.0010)	7.0313	7.0578	7.0839
$\mathbf{Dur_H}$	2.0530	(0.0012)	2.0496	2.0530	2.0559
Panel B: Conditional Belief					
	Mean	**(St.Err)**	**2.5%**	**50%**	**97.5%**
$\mathbf{E_t}[\mu_\mathbf{t+1}]$	0.2151	(0.0113)	-0.3639	0.2300	0.6552
$\mathbf{Std_t}[\mu_\mathbf{t+1}]$	0.0312	(0.0160)	0.0135	0.0261	0.0662

Figure A.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

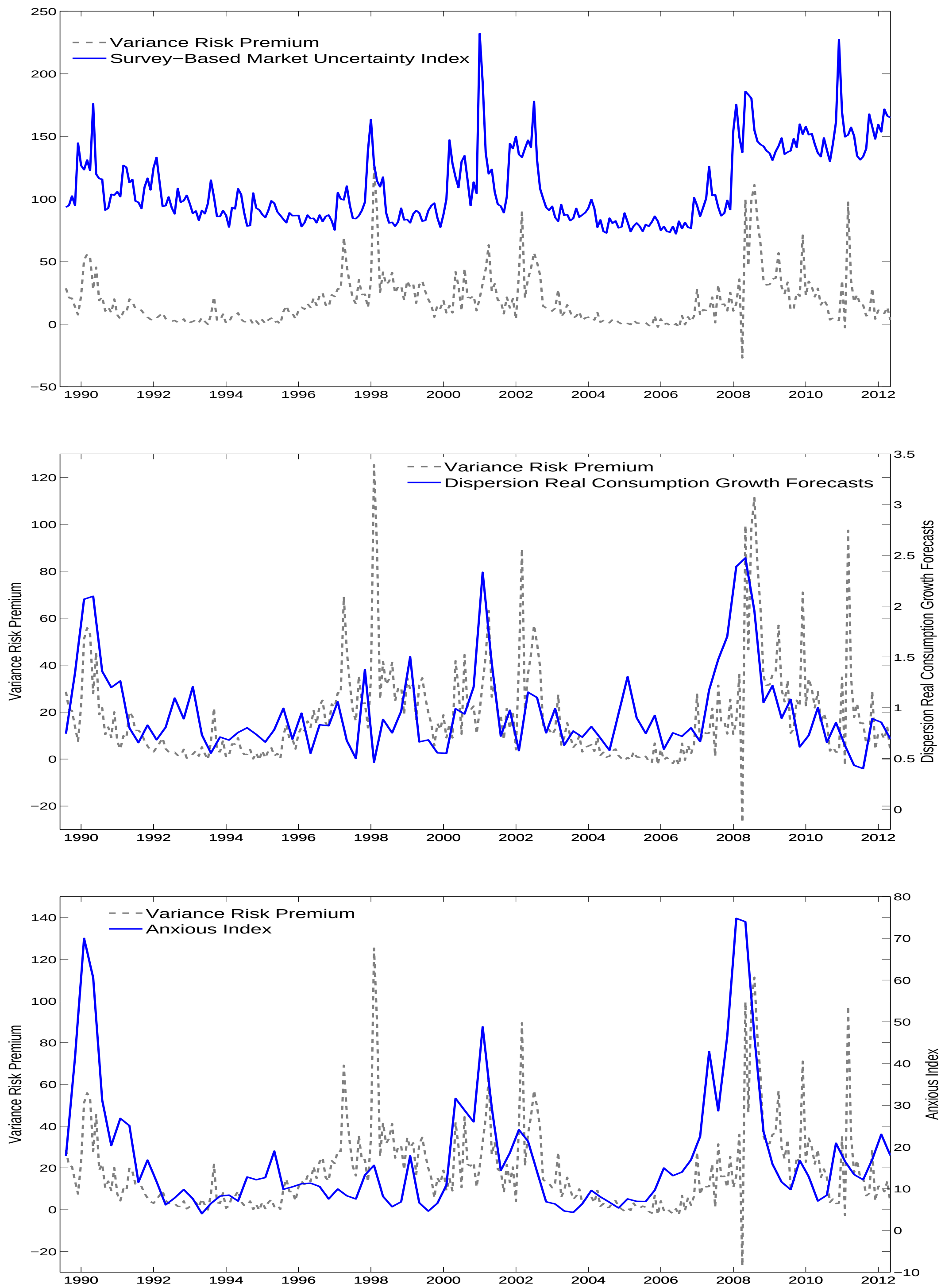


Figure A.2: Conditional Variance Real Consumption Growth

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

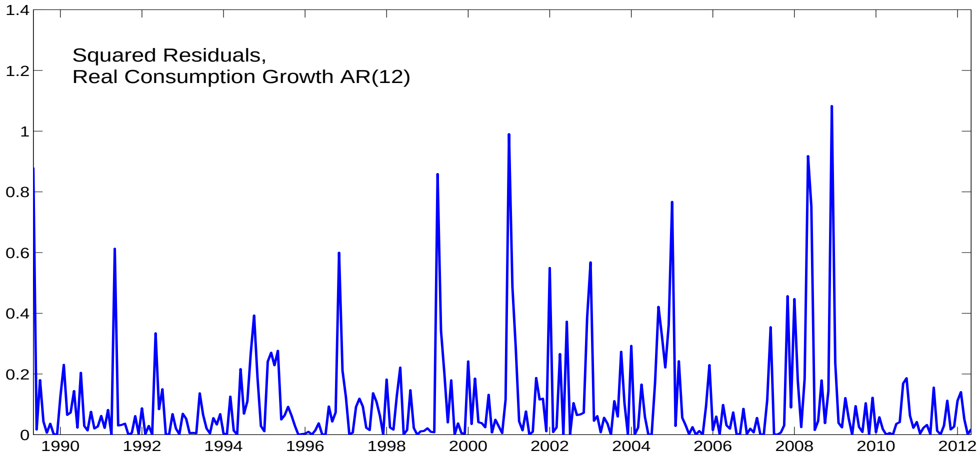
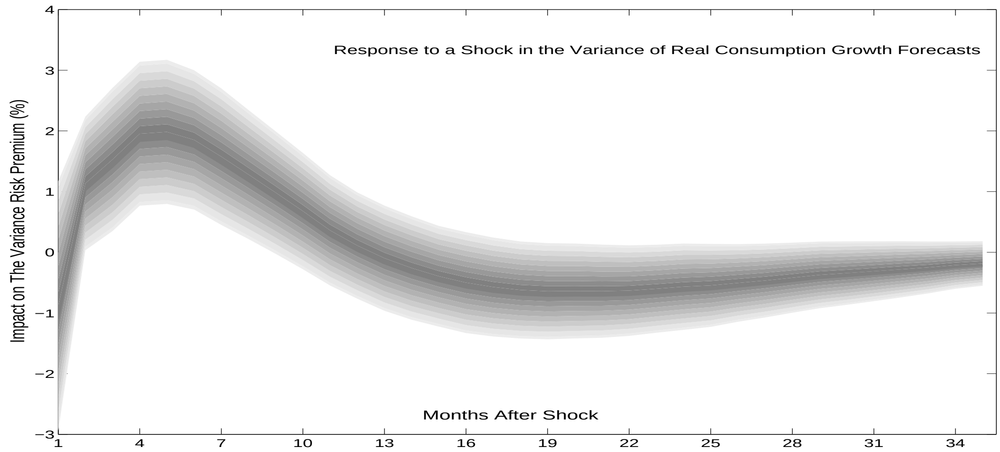
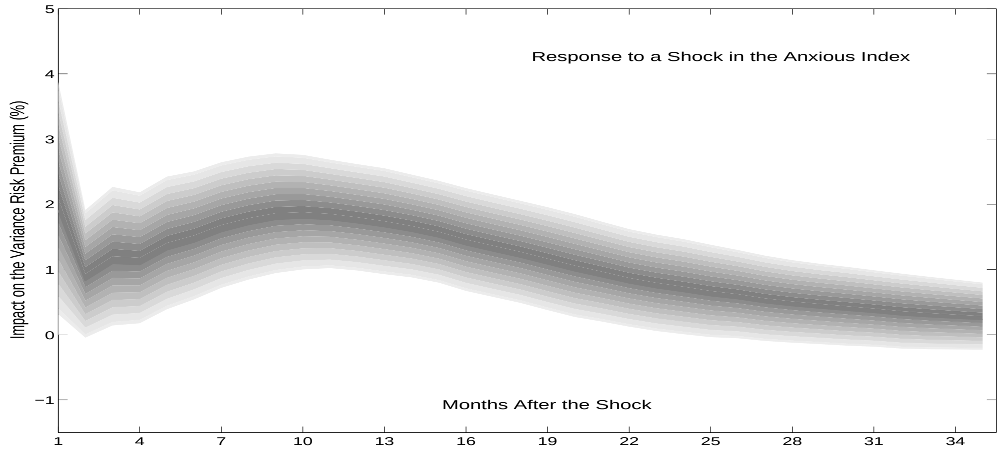


Figure A.3: Impulse Responses

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.





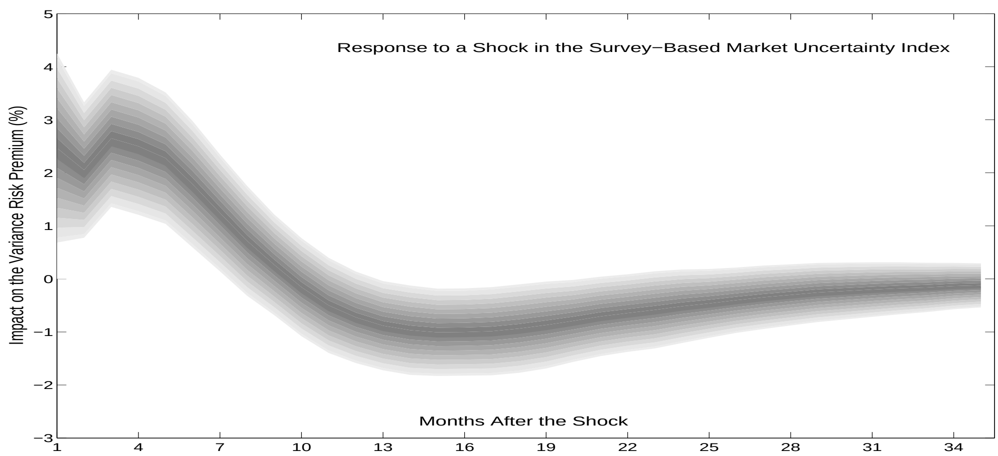
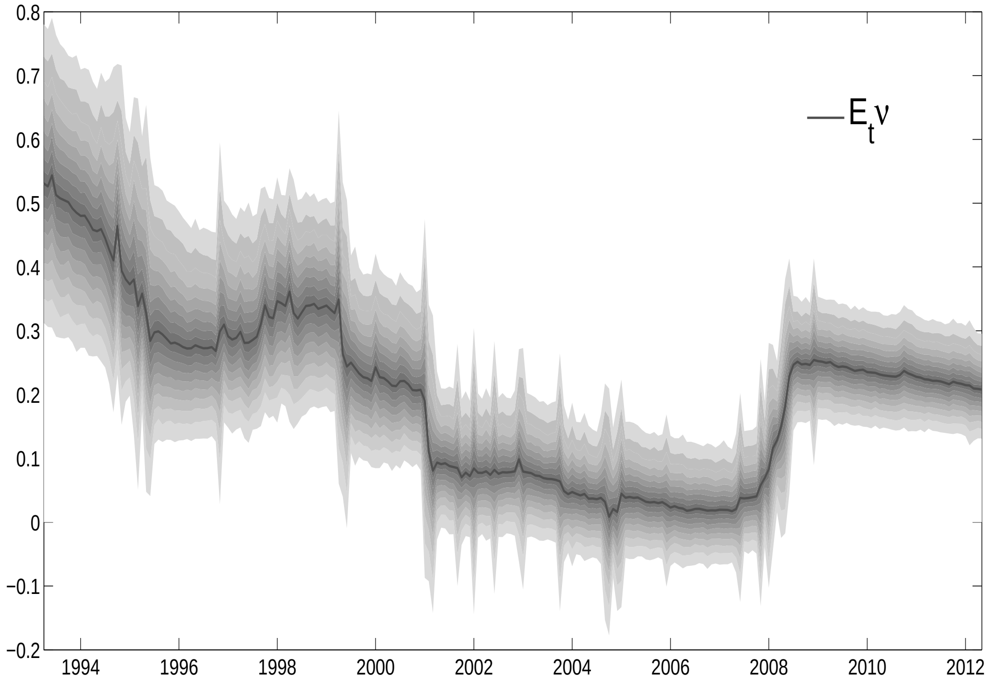


Figure A.4: Real-Time Parameters Estimates

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.



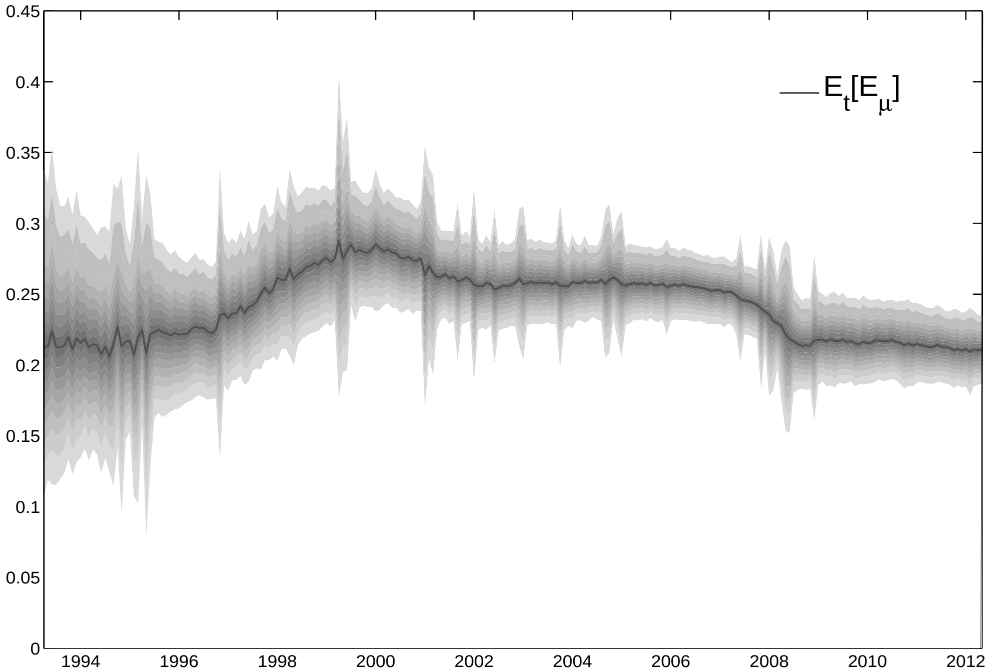


Figure A.5: Uncertainty Shocks

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.

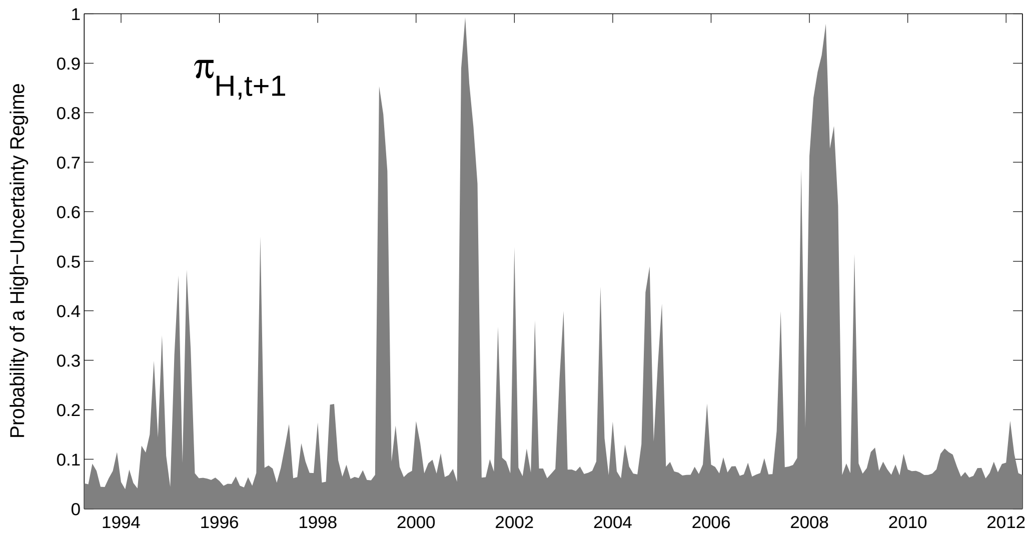


Figure A.6: Matching the Conditional Dynamics

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.

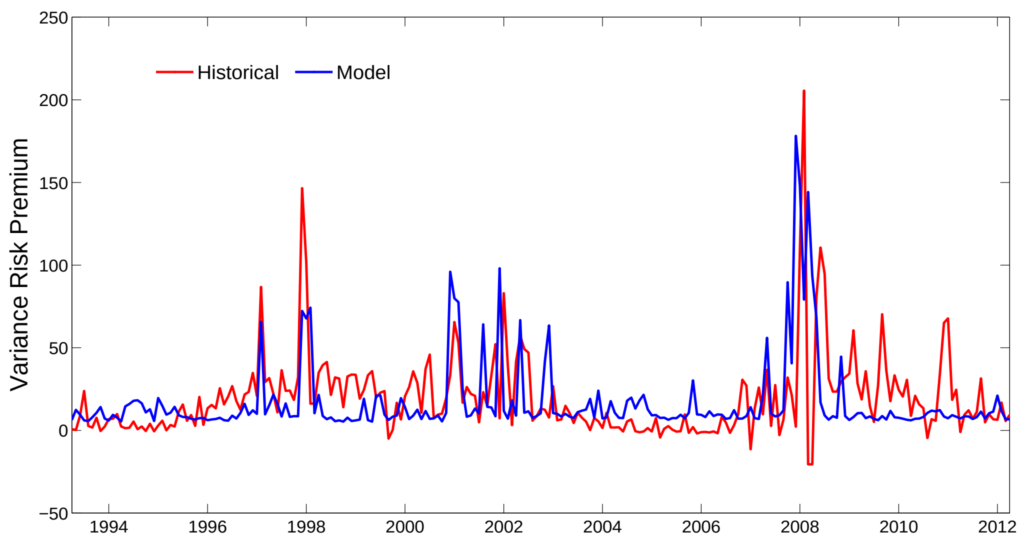


Figure A.7: The Impact of an Uncertainty Shock

This figure shows the relation between the state λ_{t+1} on both the Kalman-Gain (left-scale) and the predictive variance of the drift μ_{t+1} (right-scale).

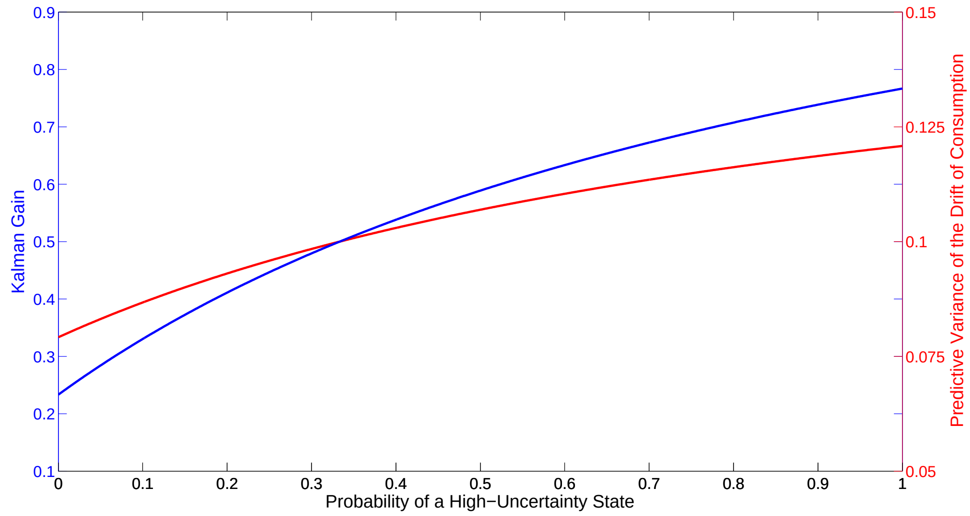


Figure A.8: The Impact of Regime Changes on the Economic Uncertainty

This figure shows the relation between the state λ_{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).

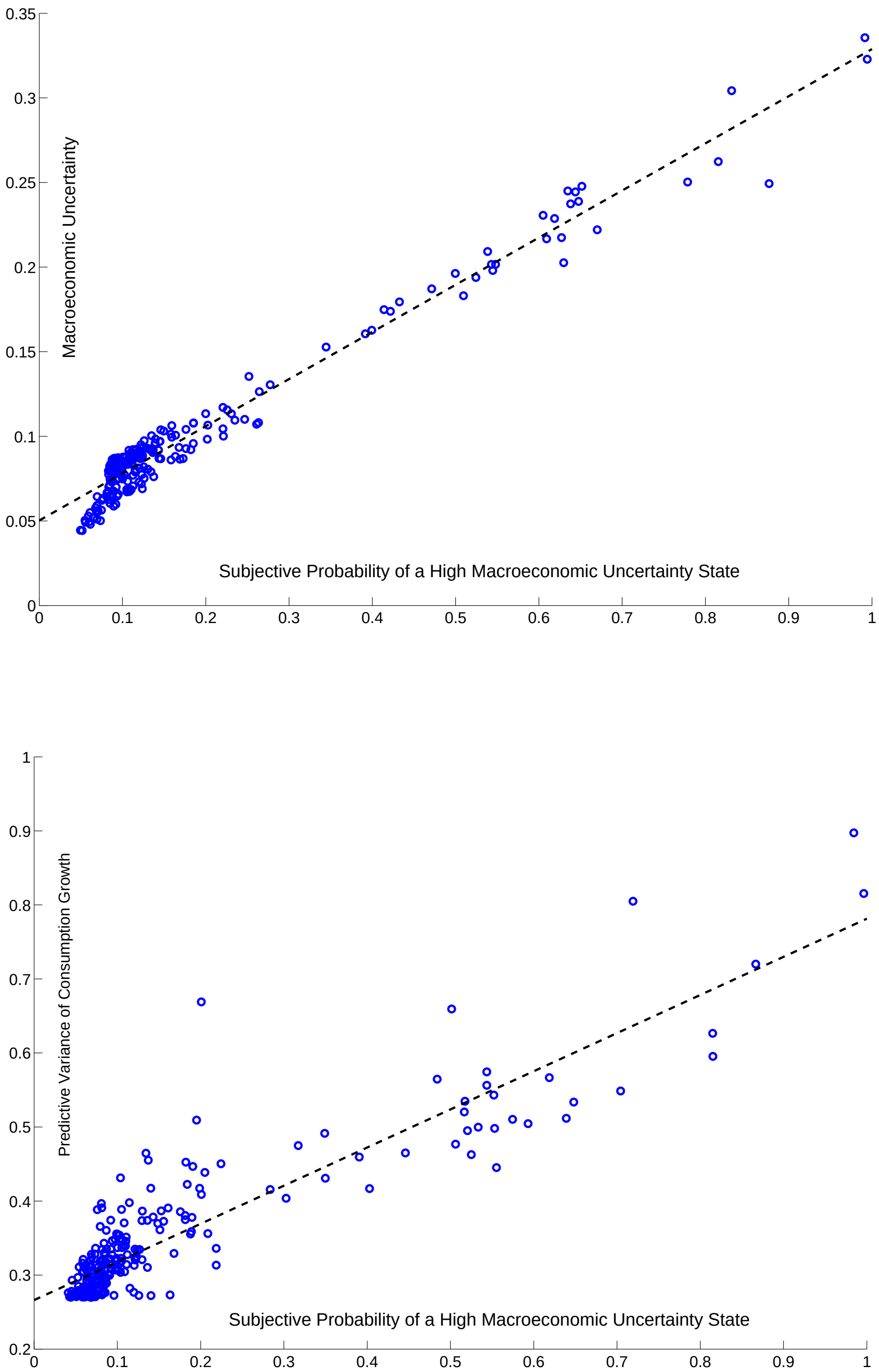


Figure A.9: The Role of Parameter Learning

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 $\psi = 3.5$.

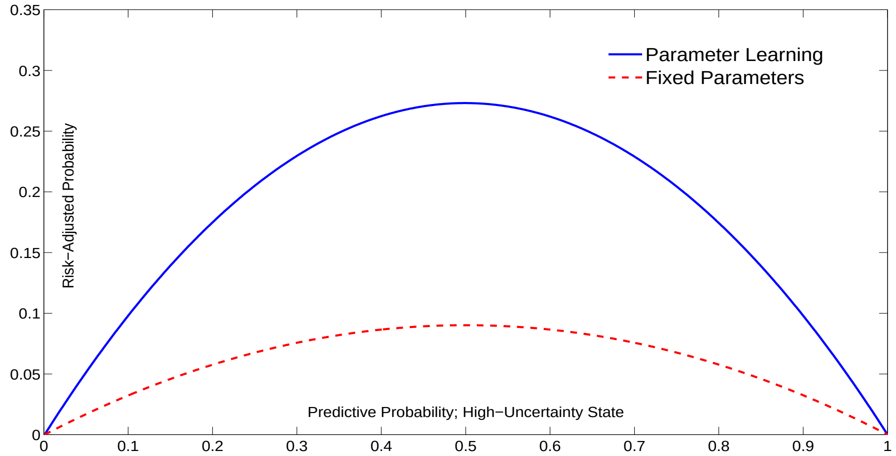


Figure A.10: Uncertainty State Probability: Structural Learning vs Rational Expectations

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.

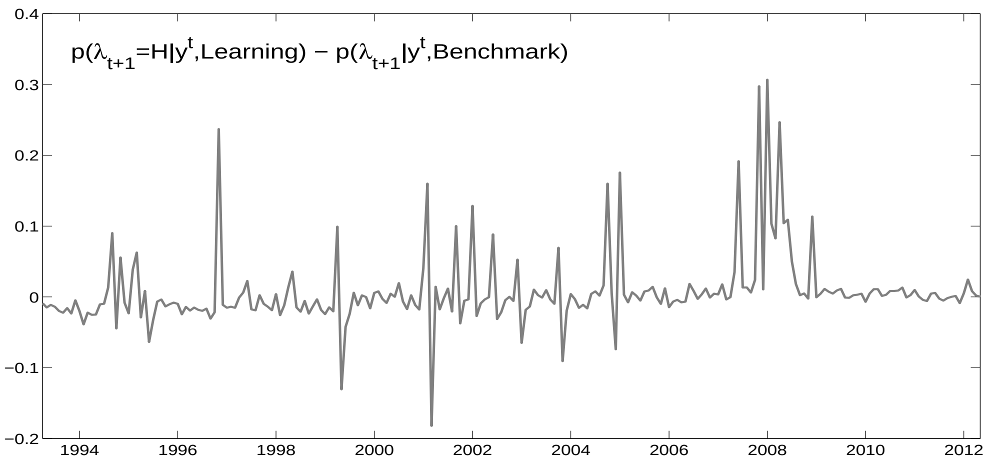


Figure A.11: The Role of Risk Aversion

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 $\psi = 3.5$.

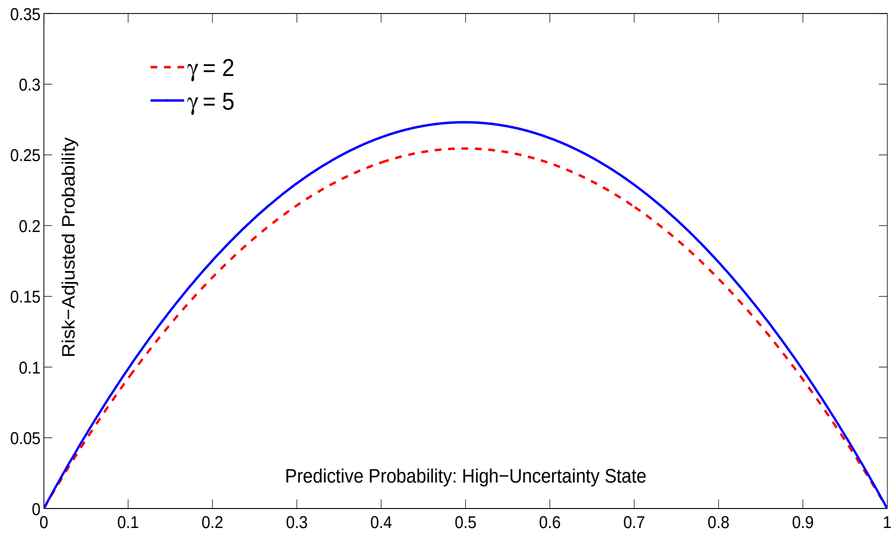
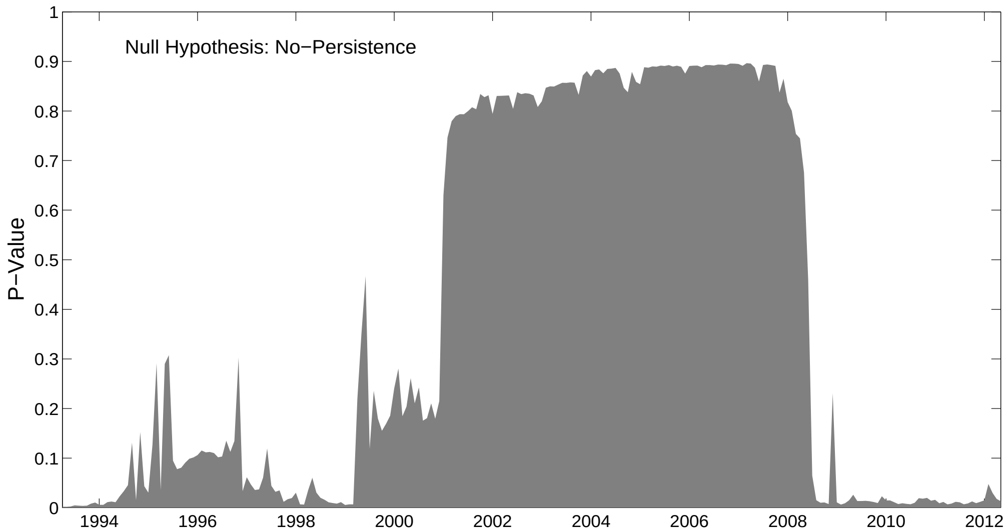


Figure A.12: Posterior Models Probabilities

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 A.13: Hypothesis Testing

This figure reports the $p(\mathcal{H}_0|y^t)$ where the null hypothesis $\mathcal{H}_0 : \nu = 0$ and $\mathcal{H}_1 : \nu \neq 0$. The sample period is 1990:01 - 2013:01, monthly. The first four years are cut as burn-in sample.



Chapter 2

Macroeconomic Factors Strike Back: A Bayesian Change-Point Model of Time-Varying Risk Exposures and Premia in the U.S. Cross-Section

with Massimo Guidolin and Francesco Ravazzolo

JEL Classification: G11, C53

2.1 Introduction

Can a selected set of macroeconomic variables explain the cross-sectional behavior of U.S. stock and bond returns? This simple question lies at the heart of the burgeoning field of macro-finance. Remarkably enough, the answer provided by at least 20 years of research on this crucial question has been predominantly negative (see e.g., Breeden, Gibbons, and Litzenberger 1989, Chan and Lakonishok 1998, McQueen and Roley 1993, Shanken and Weinstein 2006): although occasional nuances to this fundamentally negative result have been reported (e.g., Flannery and Protopapadakis 2002, using high frequency data; Kramer 1994, to explain seasonality in stock returns; Henkel and Nardari 2011, and McQueen and Roley 1993, conditioning on the state of the economy), it is common wisdom that macroeconomic factors can hardly explain the dynamics of asset valuations for U.S. stock and bond portfolios. Such a disconnect between changes in aggregate variables representing sources of systematic risk—like in the case of real output and inflation news—and asset returns has long represented a puzzle with far-reaching implications for the finance profession, because it reflects in persistent difficulties at relating the determination of the cost of capital at the firm level, performance evaluation in delegated asset management, and risk management assessments to business cycle and aggregate economic conditions. Moreover, this disconnect clashes with a body of theoretical work that shows that on average, excess returns (risk premia) should be driven by macroeconomic forces (see, e.g., Cochrane 2005).

In this chapter we propose and estimate through Bayesian methods a flexible parametric multi-factor, stochastic volatility asset pricing model in which both risk exposures (betas) and the prices of a number of macroeconomic risk factors are time-varying and effectively explain the cross-section of U.S. stock and bond returns. Time variation is modelled as a latent, change-point process. We show that an explicit parameterization of latent breakpoints in betas and risk premia plays a key role: when the model is made sufficiently flexible by capturing potential breaks in risk exposures, premia, and in the process of stochastic idiosyncratic variance, macro factors re-gain a solid explanatory and pricing power for the U.S. cross-section that is not already embedded in the market risk factor.

By comparing our baseline model with restricted versions (e.g., homoskedastic or classical random walk time-varying betas) of the same, we provide evidence that both stochastic volatility and infrequent but possibly large parameter shifts are key drivers of the capability of the model to capture cross-sectional return dynamics. Finding evidence of a precisely estimated link between time-varying betas on selected macroeconomic factors (e.g., aggregate market returns, the rate of growth of industrial production, shocks to the inflation rate, the spread between long- and short-term yields, etc.) and stock and bond excess returns speaks to the heart of finance theory, because any evidence uncovered bears on the fundamental issue of the features of the general pricing mechanism (the stochastic discount factor, SDF), underlying security prices: In our chapter we show that it is both possible and useful to connect such an SDF (assumed to exist and to be unique) to macroeconomic risk.

A related but not less important question concerns the most appropriate methods to learn about the SDF that underlies the cross-section of asset returns. Our chapter offers a contribution to the voluminous literature that has tackled both questions specializing to a particular set of linear multi-factor models, offers a novel statistical framework to implement such models, and shows how this works using an empirically relevant application. With reference to an application to 40 years of monthly data on excess returns for 23 portfolios of securities traded in the U.S., we show that while commonly used methods to estimate macro-based factor models fail to lead to sensible conclusions, an encompassing MCMC algorithm that allows for instability in factor exposures and risk premia delivers encouraging results.[1]

Following the seminal work of Fama and MacBeth (1973), two-step multi-factor asset

[1] In particular, we use a Gibbs sampler that is a special case of Markov chain Monte Carlo methods. Johannes and Polson (2005) survey their applications in finance.

pricing models (MFAPMs) have been commonly used to estimate multi-factor models. Fama-MacBeth's (henceforth F-MB) approach corresponds to a simple algorithm that uses a first set of rolling window, time series regressions to obtain (often, least-square) estimates of the risk exposures, followed by a second-pass set of cross-sectional (across assets or portfolios) regressions that using the first-pass rolling window betas as inputs to derive time-varying estimates of the associated risk premia. The problems with this methodology are notorious (see e.g., the discussions in Geweke and Zhou 1996; Jagannathan and Wang 1998; Zhou 1999):[2] most inferential statements made as a result of the second-pass would be valid if and only if one could assume that the first-pass betas were fixed in repeated samples, which clearly clashes with them being least squares estimates and therefore random sample statistics. Unless additional and strong assumptions are introduced, this creates a serious problem with generated regressors being used in the second-step, which tends to make invalid most of the inferential statements commonly made when the resulting errors-in-variables problems are ignored (see e.g., Pagan 1984; Geweke and Zhou 1996; Shanken 1992).[3] F-MB's approach also suffers from another problem: although identifying time-variation in risk exposures and risk premia with a rolling window least square estimation is robust because it is nonparametric, the length of the window is usually chosen in an arbitrary way and this can result in severe efficiency loss (see e.g., Maheu and McCurdy 2009; Shanken and Weinstein 2006).

To overcome these problems, we introduce a different approach where time variation in risk exposures and premia is explicitly modelled as a break-point process.[4] Specifically,

[2]Our chapter is not about how to produce better (valid) standard errors than under F-MB's methods in tests involving panels of data, to take into account cross- and own-serial correlation effects. Geweke and Zhou (1996) discuss the difference between the two endevours.

[3]In practice, the classical two-step procedure either does not provide a known asymptotic distribution for functions of interest in applied work or these asymptotic distributions may not be reliable in finite samples, even when they are available.

[4]A Bayesian approach is not the only solution to the issues with the F-MB approach. For instance, under the assumption that, given the realization of factors, asset returns have a conditional joint distribution with constant covariance matrix, Shanken (1992) shows that the F-MB standard errors overstate the precision of the estimated parameters and derives appropriate adjustments. Burmeister and McElroy (1988) employ a restricted nonlinear multivariate SUR model.

we model risk exposures as latent stochastic processes in a mixture innovation framework as in Giordani and Kohn (2008); also Maheu and Gordon (2008) stress the importance of incorporating breaks when modelling asset returns. The parameters of interest are constant unless a break-point variable takes a unit value, in which case the parameters are allowed to jump to a new level, as a result of a normally distributed shock. Furthermore, to consistently overcome the problems with generated regressors, the model is estimated in a single step by using a Bayesian approach, following the seminal work by Geweke and Zhou (1996), Jostova and Philipov (2005), McCulloch and Rossi (1990), and Nardari and Scruggs (2007). Like Geweke and Zhou, we provide an exact finite sample statistical framework for testing multi-factor models. By construction, our approach represents a single-step procedure that yields exact inferences; given the fact that there are unobservable factors in the return generating process, our framework implicitly incorporates this uncertainty into our inferences. Moreover, our approach makes it possible to compute the posterior distribution of virtually any function of the parameters that can be useful to implement economic tests (e.g., variance ratio and decomposition tests in Section 5). However, differently from Geweke and Zhou (1996) and Nardari and Scruggs (2007), we deal with the increasingly realistic case of time-varying risk exposures and prices and consider the economically important case of pre-determined macroeconomic factors instead of extracting them as latent principal components.

Our main results are as follows. First, using a variety of metrics—such as Bayes factors and average pricing errors—we obtain evidence of the importance of capturing instability *both* in betas and in volatility; additionally, simpler time-varying parameter models in which betas follow random walk processes subject to breaks that are frequent but of modest size appear to be outperformed by a change-point model. The Bayesian (posterior median) estimates of the risk premia are stable and a few of them are precisely estimated. Moreover, variance decompositions show that by considering model instability together with parameter uncertainty, the amount of information explained by the factor model increases with respect not only to a standard F-MB, but also vs. specific restrictions on the dynamics of either the factor sensitivities or idiosyncratic variance. Second, the Bayesian time-varying betas, stochastic volatility model leads to economically plausible estimates with reference to an application for which the two-step approach fails to provide plausible insights and would lead to a MFAPM rejection. For instance, a two-step F-MB approach leads to all the 23 test portfolios displaying large and persistent mis-pricing during our sample period. On the contrary, in the Bayesian case, the values of the posterior medians of the same parameters indicate the absence of large mis-pricings. Third, best fitting mixture model reveals that a number of macroeconomic factors significantly contribute to explain the U.S. cross-section of stocks and bonds, and that the associated time-varying unit risk premia are accurately estimated in our sample and have sensible signs and magnitudes. Fourth, the F-MB approach shows that idiosyncratic risk is large for most portfolios investigated and highly unstable, with detectable trends; in our Bayesian model, when all the uncertainty is taken into account, there is no longer strong evidence of trends in idiosyncratic risk.

2.2 The Pricing Framework

Our empirical work is based on model from the multi-factor linear class introduced, among others, by Ferson and Harvey (1991). Multi-factor asset pricing models (MFAPMs) posit a linear relationship between asset returns and a set of systematic factors that are assumed to capture business cycle effects on beliefs and/or preferences (as summarized by a SDF with time-varying properties, see e.g., Cochrane 2005) and hence on risk premia. If we call the process for the risk factors $F_{j,t}$ $(j = 1, ..., K)$ and $r_{i,t}$ the *excess* return on asset or portfolio $i = 1, ..., N$, a typical MFAPM can be written as:

$$r_{i,t} = \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \epsilon_{i,t} \qquad \epsilon_{i,t} \sim (0, \sigma^2_{i,t}). \tag{2.1}$$

It is customary to assume that $E[\epsilon_{i,t}] = E[\epsilon_{i,t}F_{j,t}] = 0$ for all $i = 1, ..., N$ and $j = 1, ..., K$.[5] The advantage of MFAPMs such as (2.1) consists of the fact that a number of systematic risk factors $K << N$ may efficiently capture relatively large portions of the variability in asset returns. Importantly, even though the notation $\beta_{ij,t}$ emphasizes that the factor loadings are allowed to be time-varying, such patterns of time variation are in general left unspecified. Finally, the $\beta_{i0,t}$ coefficients are often interpreted as abnormal returns on asset i "left on the table" after all risks ($F_{j,t}$, $j = 1, ..., K$) and risk exposures ($\beta_{ij,t}$, $j = 1, ..., K$) have been taken into account.

In the conditional version the APT or of the intertemporal CAPM (ICAPM), the expected excess return on asset i (i.e., the risk premium) may be approximately related by an arbitrage argument to its "betas" and the associated unit risk premia,

$$E[r_{i,t}|\mathbf{Z}_{t-1}] \simeq \lambda_0(\mathbf{Z}_{t-1}) + \sum_{j=1}^{K} \beta_{ij,t|t-1}\lambda_j(\mathbf{Z}_{t-1}), \tag{2.2}$$

where both the betas and the risk premia are conditional on the information publicly available at time t, here summarized by the $M \times 1$ vector of "instruments" $\mathbf{Z}_t$ in the information set, that capture effects of the state of the economy on the risk premia. The framework in (2.1)-(2.2) just describes a general conditional pricing framework that is known to hold under a variety of alternative assumptions.[6] However, a variety of alternative methodologies have been proposed to perform three related tasks which affect the empirical performance of (2.1)-(2.2):

(i) how many factors ought to be selected, i.e., picking an appropriate value for K;

(ii) given K, devising a methodology to be able to rank competing factors;

[5]The time-varying process for idiosynchratic risk, $\sigma^2_{i,t}$, is left unspecified by asset pricing theory and can be thought of as one of the standard frameworks popular in empirical finance, such as a simple GARCH(1,1) or a stochastic volatility model.

[6]Cochrane (2005) shows that (2.1)- (2.2) holds when the stochastic discount factor can be written as an exact linear function of the systematic risk factors F_1, F_2, ..., F_K. Under alternative conditions, (2.1) may hold asymptotically, as $N \to \infty$.

(iii) estimating the factor loadings $\{\beta_{ij,t}\}$ (over time and for each possible pair i, j) and the risk premia λ_{jt} (over time and for each possible j).

These tasks are logically distinct from the formulation of the framework and they have a statistical nature. In this chapter we align ourselves to a number of chapters in the empirical finance literature (see e.g., Burmeister and McElroy 1988; Chan and Hsieh 1985; Chen and Ross 1986; Kramer 1994) as far (i)-(ii) are concerned—which means that we pre-select both K and which specific macroeconomic risk factors ought to be considered in the light of the existing literature—and provide an alternative, arguably more flexible econometric approach to accomplish task (iii).

2.2.1 The Standard Two-Stage Approach

The standard approach is the classical, two-stage procedure *à* la Fama and MacBeth (1973) also used by Ferson and Harvey (1991) and very popular in empirical finance: In the first step, for each of the assets, the factor betas in (2.1) are estimated via a simple rolling window OLS. Once estimates of the risk exposures, i.e., the betas, are obtained, these are plugged in a second step as covariates testing the equilibrium restriction in (2.2) (see Appendix A or Shanken and Weinstein (2006), for details). To favor comparability between our Bayesian implementation with stochastic volatility and the standard two-step approach, all the results in this chapter are obtained under the assumption that idiosyncratic variance, $\sigma^2_{i,t}$, follows a standard GARCH(1,1) process, $\sigma^2_{i,t} = \gamma_{i0,t} + \gamma_{i1,t}\epsilon^2_{i,t-1} + \gamma_{i2,t}\sigma^2_{i,t-1}$.

Although widely used, the two-stage F-MB approach has a number of well-known statistical drawbacks. First, the second stage regression used to test for the equilibrium restriction (2.2) suffers from generated regressor (error-in-measurement) problems as the estimated first-stage, rolling window beta estimates $\hat{\beta}^{60}_{ij,t-1}$ are used as regressors on the right-hand side. For instance, Ang and Chen (2007) argue that when the cross-sectional estimates of the betas $\hat{\beta}^{60}_{ij,t-1}$ co-vary with the underlying but unknown risk premia, F-MB yields biased and inconsistent estimates of the risk premia themselves. Unfortunately, this co-variation is extremely likely: for instance, the view that during business cycle downturns both the quantity of risk and the unit risk prices increase seems to be widely held, because recessions are characterized by higher systematic uncertainty as well as by lower "risk appetite" (e.g., Campbell and Cochrane, 1999). Second, for instance as emphasized by Jostova and Philipov (2005) with reference to a single-factor conditional CAPM, when parameters in linear asset pricing models are estimated from the data, their uncertainties should be taken into account. Third, the need to perform the necessary estimation to implement (2.1)-(2.2) in two distinct stages that use rolling windows to capture parameter instability is not only *ad hoc* but also largely inefficient because the lack of more specific parametric forms makes testing for time-variation very hard and dependent on hard-to-justify choices of the rolling window length, the updating rules applied to select whether constant or decaying weights should be applied, etc. (see the discussion in Shanken and Weinstein 2006). Yet, it is also well-known that using a rolling window approach implies that (it is optimal when) there is probability one of a break in the process of parameters on every period in which the data are sampled. This seems to represent a rather extreme and often implicit assumption concerning the amount of instability represented in the data (see Maheu and McCurdy 2009).

2.2.2 Traded vs. Non-Traded Factors

One problem with (2.1) is the difficulty of interpreting $\beta_{i0,t}$ ("Jensen's alpha") when some of the risk factors are not traded. Although analyses that use (2.1) to decompose excess returns may be informative, unless all the factors are tradable portfolios, it is impossible to interpret any non-zero β_{i0} as an abnormal return on asset i: there reason is that there may be a considerable difference between the theoretical alphas from an estimated model, and the actual alpha that an investor may harvest from by trading assets on the basis of the MFAPM. To eliminate such a possibility, we follow the literature (see e.g., Chan and Lakonishok 1998; Lamont 2001) and proceed as follows. When an economic risk factor is already measured in the form of a return (e.g., this is the case of the U.S. market portfolio, real T-bill rates, the liquidity and bond factors, term structure and default spreads), we directly use the associated returns as a mimicking portfolio. When a factor is not itself an (excess) return (this is the case of macroeconomic variables such as industrial production growth, unexpected inflation, and real consumption growth), we construct the corresponding $K' \leq K$ mimicking portfolios by projecting the non-traded factors onto the space of excess returns of base assets and a set of predictive variables ($j = 1, ..., K'$):

$$F_{j,t} = a_j + \mathbf{b}_j'\mathbf{x}_t + \mathbf{c}_j'\mathbf{z}_{t-1} + \varepsilon_{j,t} \qquad \varepsilon_{j,t} \text{ IID } (0, 1), \tag{2.3}$$

where $\mathbf{x}_t$ is a vector of excess returns on the base assets (in this case, all defined to be zero investment portfolios) and $\mathbf{z}_{t-1}$ denotes a vector of instruments that have the ability to predict returns. The resulting returns on the ith factor mimicking portfolio (FMP henceforth) is then defined as $FMP_{j,t} = \hat{a}_j + \hat{\mathbf{b}}_j'\mathbf{x}_t$, and collects the fitted component of a factor that is unpredictable on basis of past information and that at the same time may be replicated by trading base assets using weights $\hat{\mathbf{b}}_j$. Note that the coefficients a_j and $\mathbf{b}_j$ do not need to add up to one because the base assets are zero-investment portfolios (see Lamont 2001). The base assets include six equity zero investment portfolios with different book-to-market and size characteristics as well as the returns on long-term minus the returns on the short term government bonds and the return on long-term corporate minus the return on long-term government bonds. We choose these assets for their ability to span large "portions" of the return space. The instruments are the lagged yield spread of long-term Treasury vs T-bills, the lagged yield spread of long-term corporate minus the yield on long-term government bonds, and the lagged real short-term rate.

2.3 A Bayesian State-Space Approach

Our discussion of the standard F-MB two-step procedure implies that we need to: (1) avoid using estimates of the first-stage betas as if these were observed variables that may be constant in repeated samples; (2) fully account for parameter uncertainty; and (3) make an effort to produce a sensible model of parametric instability—here in the form of structural breaks—to reflect the commonly perceived —see e.g., Ghysels (1998), Kramer (1994) and references therein—fact that both the relationship between excess returns and factors, namely risk exposures ($\beta_{ij,t}$), the risk premia (λ_j, for $i = 1, ..., N$ and $j = 1, ..., K$), and possibly also residual idiosyncratic variances ($\sigma^2_{i,t}$) stochastically change over time.[7] We therefore develop a new Bayesian estimation approach in which:

- The error in measurement due to the stochastic nature of the betas is avoided following Geweke and Zhou (1996), McCulloch and Rossi (1990), and Nardari and Scruggs (2007), by characterizing the joint posterior distribution of betas and risk premia such that both states and parameters are jointly estimated in a single step.

- Parameter uncertainty is fully addressed by using Bayesian techniques that integrate the joint posterior to find the joint predictive density of the variables of interest (see e.g., Tu 2010).

- Model instability is captured by introducing stochastic breaks in the dynamics of the factor loadings as well as of idiosyncratic volatility (see e.g., Tu 2010).

Specifically, we characterize the relationship between excess returns and factors and the time-varying dynamics in factor loadings and idiosyncratic volatility in a state-space

[7]Using the formal links between time-varying beta factor models and predictability established in Ferson and Korajczyk (1995), one could also infer a need for modelling instability also from recent evidence on the joint presence of unstable predictability and stochastic volatility in stock returns (see e.g., Johannes, Korteweg, and Polson 2013; Tu 2010).

form where the observation equation is the standard linear factor model (2.1)

$$r_{i,t} = \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \sigma_{it}\epsilon_{i,t} \tag{2.4}$$

where $\epsilon_t \equiv [\epsilon_{1,t}, \epsilon_{2,t}, ..., \epsilon_{N,t}]' \sim N(0, \mathbf{I}_N)$ and $E[\epsilon_{i,t}] = E[\epsilon_{i,t}F_{j,t}] = 0$ for all $i = 1, ..., N$ and $j = 1, ..., K$. The time varying parameters $\beta_{ij,t}$ and σ_{it} are described by the state equations

$$\beta_{ij,t} = \beta_{ij,t-1} + \kappa_{ij,t}\eta_{ij,t} \qquad j = 0, ..., K, \tag{2.5}$$

$$\ln(\sigma^2_{i,t}) = \ln(\sigma^2_{i,t-1}) + \kappa_{i\nu,t}\upsilon_{i,t} \qquad i = 1, ..., N, \tag{2.6}$$

where $\eta_{i,t} \equiv [\eta_{0,t}, \eta_{1,t},..., \eta_{K,t}, \upsilon_{i,t}]' \sim N(0, \mathbf{Q}_i)$ with $\mathbf{Q}_i \equiv diag(q^2_{i0},\ q^2_{i1},\ ...,\ q^2_{iK},\ q^2_{i\upsilon})$. Stochastic variation (breaks) in the level of both the beta coefficients and of the idiosyncratic variance σ^2_{it} are introduced and modelled through a mixture innovation approach as in Chan and Maheu (2002) and Giordani and Kohn (2008). The latent binary random variables $\kappa_{1ij,t}$ and $\kappa_{2i,t}$ capture the presence of random shifts in betas and/or idiosyncratic variance (see Miazhynskaia 1988; Miazhynskaia 2006). The random variable $\kappa_{1ij,t}$ takes then a value equal to one if a structural break for the jth factor in the equation for the ith asset at time t takes place. We assume that the structural breaks are independent among one another (i.e., across assets as well as factors) and over time with:[8]

$$\Pr\left[\kappa_{ij,t} = 1\right] = \pi_{ij} \quad \Pr\left[\kappa_{i\nu,t} = 1\right] = \pi_{i\nu} \qquad i = 1, ..., N \qquad j = 0, ..., K. \tag{2.7}$$

This specification is flexible and allows for both constant and time-varying parameters. When $\kappa_{ij,\tau} = \kappa_{i\nu,\tau} = 0$ for some $t = \tau$, then (2.5) reduces to (2.1) when the factor loadings and the quantity of idiosyncratic risk are assumed to be constant, as $\beta_{ij,\tau} = \beta_{ij,\tau-1}$ and

[8] The indepence across breaks is consistent with the spirit of a factor model and might not be necessarily restrictive. Indeed, the comovements between asset returns might be driven by the factor structure regardless of the nature of the structural breaks. In this perspective, our approach is fairly flexible because "it lets the data speak" about whether breaks across assets i and j are contemporaneous or not.

$\ln \sigma^2_{i,\tau} = \ln \sigma^2_{i,\tau-1}$. However, when $\kappa_{1ij,\tau} = 1$ and/or $\kappa_{2i,\tau} = 1$, then a break hits either a beta or idiosyncratic variance or both, and instability is then captured by the random walk dynamics $\beta_{ij,\tau} = \beta_{ij,\tau-1} + \eta_{ij,\tau}$ and $\ln(\sigma^2_{i,\tau}) = \ln(\sigma^2_{i,\tau-1}) + \upsilon_{i,t}$ (or $\sigma^2_{i,\tau} = \sigma^2_{i,\tau-1} \exp(\upsilon_{i,\tau})$). The flexibility of the specification in (2.5) stems from the fact that risk exposures, $\beta_{ij,t}$, and idiosyncratic risks, $\sigma^2_{i,t}$ (as in Nardari and Scruggs 2007), are allowed to change on every time period, but they are not imposed to be changing at every point in time. In our view, this helps to side-step the difficult task of persuading a Reader that the assumed dynamics represents the "right" kind: given our uninformative priors, if the data need frequent breaks in betas of a small size, the posterior of the corresponding parameters will provide indications in this direction; similarly, if the data need a stochastic volatility process with frequent shifts, posterior estimates will give appropriate indications, etc.

Note that because when a break affects the betas and/or the variances, the random shift is measured by variables collected in $\eta_{i,t}$, we can interpret $\mathbf{Q}_i$ as the "size" of the break: a large q^2_{ij} means for instance that whenever $\beta_{ij,t}$ is hit by a break, i.e., $\kappa_{ij,t} = 1$, such a shift is more likely to be large (in absolute value). This process for factor loadings and idiosyncratic residual risk is different and more flexible when compared to the structures typical of the time-varying parameter finance literature in which loadings are assumed to vary on every period and usually according to simplistic AR(1) structures with high persistence and small variance for the shocks, such as $\beta_{ij,t} = \phi_{1ij}\beta_{ij,t-1} + \eta_{ij,t}$ and $\ln(\sigma^2_{i,t}) = \phi_{2i} \ln(\sigma^2_{i,t-1}) + \upsilon_{i,t}$ with ϕ_{1ij} and ϕ_{2j} close to but less than one. Moreover, even though time-varying betas, $\beta_{ij,t}$, clearly depend only on the information up to time t so that the spirit of (2.2) applies, (2.5) avoids any parameterization of the dependence of the betas from the instruments in $\mathbf{Z}_t$ which is beneficial as Ghysels (1998) and Harvey (2001), among others, have noted that the estimates of factor loadings obtained from the explicit use of instrumental variables are very sensitive to the variables considered.

The cross-sectional pricing restrictions in (2.2) are characterized through the multivariate linear model

$$r_{i,t} = \lambda_{0,t} + \sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1} + e_{i,t} \qquad i = 1, ..., N, \tag{2.8}$$

where $e_{i,t} \sim N(0, \psi^2)$ and $\beta_{ij,t|t-1}$ measures the *predictive* time t sensitivity of asset i to factor j, based on all information available up to time $t-1$. Because we adopt a single-step Bayesian estimation strategy, and the unknown betas and risk premia are multiplied with one another, in practice (2.8) imposes a set of non-linear restrictions in estimation. Note that under the assumption of correct specification of the asset pricing model, the return on the zero beta portfolio $\lambda_{0,t} = 0$ or, at least, the average over time of the $\lambda_{0,t}$s ought to be zero, implying that only the assumed risk factors are explaining the risk premia on the different assets and portfolios. As clarified by Nardari and Scruggs (2007), this is one of the payoffs of our ex-ante definition of K predetermined macroeconomic factors. In our setting $\beta_{ij,t|t-1}$ represents a draw from the predictive distribution of the state dynamics in (2.5), which is obtained by integrating out both the probability of recording a structural break and the uncertainty about the size of the break itself (see Maheu and McCurdy 2009). This is the exact analog of the logic followed by Ferson and Harvey (1991) who have emphasized the importance that in the implementation of factor models the time t excess return on asset i should be determined by investors with reference only to information available up to time $t-1$. The Bayesian paradigm allows to go one step further to properly capture the forward looking nature of the asset pricing model. As in Geweke and Zhou (1996), the risk premia $\lambda_t \equiv (\lambda_{0,t}, \lambda_{1,t}, ..., \lambda_{K,t})'$ are estimated jointly with the factor loadings $\mathbf{B}_t \equiv \{\beta_{ij,t}\}_{i=1}^{N}{}_{j=0}^{K}$, the (log of the) idiosyncratic variances $\sigma_t^2 \equiv (\sigma_{1t}^2, \sigma_{2t}^2, ..., \sigma_{Nt}^2)'$, as well as the other parameters $\mathbf{\Theta} = \{\theta_i\}_{i=1}^{N}$, with $\theta_i \equiv (\mathbf{q}_i^2, \pi_i)'$, where $\mathbf{q}_i^2 \equiv (q_{i0}^2, q_{i1}^2, ..., q_{iK}^2, q_{i\nu}^2)'$ is the vector of conditional variances of the factor loadings and the idiosyncratic risks and $\pi_i \equiv (\pi_{i0}, ..., \pi_{iK}, \pi_{i\nu})'$ is the vector of break probabilities for the ith asset. The time variation in risk premia is inherited by the dynamics in portfolio sensitivities, $\{\mathbf{B}_t\}_{t=1}^{T}$. Therefore, even though the dynamics of λ_t is not specified, the instability of the betas is by construction reflected in the risk premia as well. Appendix B provides additional details on the estimation algorithm.

2.3.1 Special Cases

The model presented in (2.4)-(2.7) is the most general specification we consider in this chapter. We will call this model B-TVB-SV specification indicating that we consider a Bayesian (B), Time-Varying Betas (TVB) and Stochastic Volatility (SV) framework. Here the words time-varying and stochastic for the betas and the volatilities are synonymous of structural breaks in both risk exposures and the idiosyncratic risks. Of course, the B-TVB-SV model is richly parameterized and it cannot be ruled out that issues related to over-parameterization may arise. Moreover, many of our economic implications might be driven by details of the parameterization of the change point process in (2.4)-(2.7). Therefore, for comparative purposes, we consider a number of alternative restrictions on the dynamics of the state equation:

1. $\kappa_{i\nu,t} = 0 \; \forall i,\, t$, i.e. a constant idiosyncratic volatility model:[9]

$$r_{i,t} = \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \sigma_i \epsilon_{i,t} \qquad i = 1, ..., N,$$

$$\beta_{ij,t} = \beta_{ij,t-1} + \kappa_{ij,t}\eta_{ij,t} \qquad j = 0, ..., K, \tag{2.9}$$

   under the same distributional assumption as (2.4)-(2.8). We will call this model a Bayesian homoskedastic time-varying betas model, i.e. B-TVB.

2. $\kappa_{ij,t} = 1 \; \forall i,\, j,\, t$ and $\kappa_{i\nu,t} = 1 \; \forall i, t$, time-varying parameters (TVP) according to random walk specifications common to the applied econometrics literature (Koop

[9]Trivially, the symmetric case of $\kappa_{ij,t} = \kappa_{iv,t} = 0 \; \forall t$ implies that $\beta_{ij,t} = \beta_{ij,t-1} = \beta_{ij}$ and $\ln(\sigma^2_{i,t}) = \ln(\sigma^2_{i,t-1}) = \ln(\sigma^2_i)$ and consists of the two-step Fama-MacBeth model with constant betas and idiosyncratic variances.

and Potter 2007). We implement a Bayesian time-varying parameter model (B-TVP) considered in Jostova and Philipov (2005):

$$r_{i,t} = \beta_{i0,t} + \sum_{j=1}^{K} \beta_{ij,t} F_{j,t} + \sigma_{it}\epsilon_{i,t}$$

$$\beta_{ij,t} = \beta_{ij,t-1} + \eta_{ij,t} \quad j = 0, ..., K, \qquad \ln(\sigma_{i,t}^2) = \ln(\sigma_{i,t-1}^2) + \upsilon_{i,t} \quad i = 1, ..., N. \qquad (2.10)$$

The B-TVP-SV (shortened to B-TVP) model assumes a unit probability of breaks (even though this are of a small size) in the dynamics of the states $\beta_{ij,t}$ and $\sigma_{i,t}^2$ at each point. This is a strict assumption which is not necessarily supported by the data as we will document below.

The constant volatility B-TVB specification is used to highlight the effects of instabilities in residual variances. The B-TVP is used as a competing specification to show the benefit of considering the more parsimonious, occasional large breaks in (2.5)-(2.7) as opposed to small, frequent (continuous) breaks.

2.3.2 Prior Specification

We estimate (2.5) using a Bayesian approach. Presumably, such an approach represents the only numerically feasible estimation method for a model with the features of the B-TVB-SV framework (see also Geweke and Zhou 1996; Jostova and Philipov 2005). Yet, it allows us to incorporate parameter uncertainty when estimating both the beta exposures and the equilibrium risk premia (see Tu 2010). This is particularly relevant because this implies that we can characterize the posterior probabilities for the unobserved binary states $\kappa_{ij,t}$ and $\kappa_{i\nu,t}$ for $t = 1, ..., T$. These can then be used to incorporate uncertainty regarding the timing of the structural breaks in the joint posterior of the state dynamics. The parameters of the model (2.4)-(2.8) are $\mathbf{\Theta} \equiv \{\theta_i\}_{i=1}^N$ with $\theta_i \equiv (\mathbf{q}_i^2, \pi_i)$, plus the risk premia λ_t which are estimated at each time t conditional on the factor exposures sensitivities according to (2.8).

For the Bayesian algorithm illustrated in Appendix B to work, we also need to specify the prior distributions for each of the model parameters. We choose a conjugate prior structure to keep the numerical analysis as simple as possible. However, we aim at minimizing the impact of the priors on the posterior estimates and on model comparisons. Accordingly, we employ relatively uninformative (imprecise) priors. As a result, the posterior densities of the parameters are driven primarily by the sample data. Furthermore, we assume that the all priors are independent. As far as the structural break probabilities are concerned, we assume a set of simple Beta distributions: $\pi_{ij} \sim Beta(a_{ij}, b_{ij})$, $\pi_{i\nu} \sim Beta(a_{i\nu}, b_{i\nu})$ for $i = 1, ...N$, $j = 1, ...K$. The parameters a_{ij}, b_{ij} and $a_{i\nu}, b_{i\nu}$ represent the shape hyperparameters and can be set according to our prior beliefs about the occurrence of structural breaks in $\beta_{ij,t}$ and $\ln(\sigma^2_{i,t})$, respectively.

For the conditional variance parameters $\mathbf{q}_i^2$, which reflect our prior beliefs about the size of the structural breaks we assume an inverted Gamma prior, $q_{ij}^2 \sim IG(\gamma_{ij}, \delta_{ij})$, $q_{i\nu}^2 \sim IG(\gamma_{i\nu}, \delta_{i\nu})$ for $i = 1, ...N$, $j = 1, ...K$ where $\gamma_{ij} > 0$, $\gamma_{i\nu} > 0$ and $\delta_{ij} > 2$, $\delta_{i\nu} > 2$ are the scale and degrees of freedom parameters, respectively, for the factor loadings and the (log-) variances. Finally, the prior distribution for the risk premia λ_t is a characterized as a standard multivariate normal distribution with independent priors, $\lambda_t \sim N_K(\underline{\lambda}, \underline{\mathbf{V}})$, $\psi^2 \sim IG(\psi_0, \Psi_0)$ for $t = 1, ..., T$. The parameters $\underline{\lambda}$ and $\underline{\mathbf{V}}$ represent the location vector and the scale matrix for the K-dimensional multivariate normal distribution; ψ_0 and Ψ_0 are the scale and degrees of freedom of the conditional variance parameters, respectively, in (2.8). Because these priors are independent of one another, the density of the joint prior distribution $p(\mathbf{\Theta})$ is given by their product. The choice of the values for the hyperparameters of the priors is discussed in detail in Appendix B.

2.3.3 Posterior Simulation

Posteriors are computed through the Gibbs sampler algorithm developed in Geman and Geman (1984) in combination with the data augmentation technique by, e.g., Frühwirth-Schnatter (1994). The latent variables $\beta_{ij,t}$, σ_{it}^2 and $\kappa_{ij,t}$, $\kappa_{i\nu,t}$ for each of the $i = 1, ..., N$ assets, each of the $j = 1, ..., K$ factors and at each time $t = 1, ..., T$, are simulated alongside the model parameters θ_i and the premia λ_t. One can think of the latent variables as nuisance parameters that are "integrated out" by the Gibbs sampler. However, to apply the Gibbs sampler we need to write down the complete likelihood function, the joint density of data and state variables. Defining $\theta \equiv \{\theta_i\}_{i=1}^N$, $\mathbf{B}_t \equiv \{\beta_{it}\}_{i=1}^T$, $\mathbf{B} \equiv \{\mathbf{B}_t\}_{t=1}^T$, $\mathbf{R} \equiv \{r_{it}\}_{i=1}^T{}_{t=1}^N$, $\mathbf{F} \equiv \{\boldsymbol{F}_t\}_{t=1}^T$, $\lambda \equiv \{\lambda_t\}_{t=1}^T$, $\mathcal{K}_\beta \equiv \{\kappa_{ij,t}\}_{j=1}^K{}_{i=1}^N{}_{t=1}^T$, $\mathcal{K}_\sigma \equiv \{\kappa_{iv,t}\}_{i=1}^N{}_{t=1}^T$, $\mathbf{\Sigma} = \{\sigma_{it}^2\}_{i=1}^N{}_{t=1}^T$, the likelihood is

$$p(\mathbf{R}, \mathbf{B}, \mathcal{K}, \mathbf{\Sigma}, \lambda|\theta, \mathbf{F}) = \prod_{t=1}^T \left\{ \prod_{i=1}^N p(r_{it}|\mathbf{F}_t, \beta_{it}, \sigma_{it}^2) p(\sigma_{it}^2|\sigma_{it-1}^2, \kappa_{i\nu,t}, q_{iv}^2) \pi_{i\nu}^{\kappa_{i\nu t}} (1 - \pi_{i\nu})^{1-\kappa_{i\nu t}} \times \right. \tag{2.11}$$

$$\left. \times [\prod_{j=0}^K p(\beta_{ij,t}|\beta_{ij,t-1}, \kappa_{ij,t}, q_{ij}^2) \times \pi_{ij}^{\kappa_{ijt}} (1 - \pi_{ij})^{1-\kappa_{ijt}}] p\left(\lambda_t, \psi^2 | \mathbf{B}_t, \mathbf{R}_t\right) \right\},$$

where $\mathcal{K} \equiv (\mathcal{K}_\beta, \mathcal{K}_\sigma)$ and $\boldsymbol{F}_t = (F_{1,t}, F_{2,t}, ..., F_{K,t})'$. Combining the prior specifications with the complete likelihood, we obtain the posterior density $p(\theta, \mathbf{B}, \mathcal{K}, \mathbf{\Sigma}, \lambda|\mathbf{R}, \mathbf{F}) \propto p(\boldsymbol{\theta}) p(\mathbf{R}, \mathbf{B}, \mathcal{K}, \mathbf{\Sigma}, \lambda|\theta, \mathbf{F})$. Our Gibbs sampler is a combination of the Forward Filtering Backward Sampling (FFBS henceforth) of Carter and Kohn (1994), and the efficient algorithm for random breaks in Gerlach and Kohn (2000). At each iteration of the sampler, we sequentially cycle through the steps:

1. Draw $\mathcal{K}_\beta$ conditional on $\mathbf{\Sigma}, \mathcal{K}_\sigma, \theta$, $\mathbf{R}$ and $\mathbf{F}$.

2. Draw $\mathbf{B}$ conditional on $\mathbf{\Sigma}, \mathcal{K}, \theta$, $\mathbf{R}$ and $\mathbf{F}$.

3. Draw $\mathcal{K}_\sigma$ conditional on $\mathbf{B}, \mathcal{K}_\beta, \theta$, $\mathbf{R}$ and $\mathbf{F}$.

4. Draw $\mathbf{R}$ conditional on $\mathbf{B}, \mathcal{K}, \theta$, $\mathbf{R}$ and $\mathbf{F}$.

5. Draw $\boldsymbol{\lambda}$ conditional on $\mathbf{B}, \mathcal{K}, \theta$, $\mathbf{R}$ and $\boldsymbol{\Sigma}$.

6. Draw θ conditional on $\mathbf{B}, \mathcal{K}$, $\mathbf{R}$ and $\mathbf{F}$.

We use a burn-in period of 1,000 and draw 5,000 observations storing every other of them to simulate the posterior distribution of parameters and latent variables. The resulting autocorrelations of the draws are very low. Appendix B provides additional details.

2.4 Empirical Results

2.4.1 Data and Descriptive Statistics

We consider a typical cross-sectional application based on a large number (23) of monthly time series sampled over the period 1972:01-2011:12 (see e.g., Mei 1993; Shanken and Weinstein 2006). The starting date is due to the availability of the complete set of instruments and corporate bond return data. The initial ten years are used to empirically elicit the priors. Our empirical analysis is implemented over the remaining 30 years of data. The series belong to two main categories. The first group, "Portfolio Returns", includes stocks, U.S. Treasuries and notes, and corporate bonds, all organized in portfolios. The stocks are listed on the NYSE, AMEX and Nasdaq (from CRSP) and sorted according to two criteria, similarly to Geweke and Zhou (1996). First, 10 industry portfolios are obtained by sorting firms according to their four-digit SIC code. Second, 10 additional portfolios are derived by sorting (at the end of every year, and recursively updating this sorting in every year in our sample period) stocks according to their size, as measured by the aggregate market value of the company's equity. All the portfolios are value-weighted. Data on long- (10-year) and medium-term (5-year) government bond returns are from Ibbotson and available from CRSP. Data on 1-month T-bill, 10-year and 5-year government bond yields and returns are from FREDII at the Federal Reserve Bank of St. Louis and from the CRSP return tapes. Data on "junk" bond returns are approximated from Moody's (10-to-20 year maturity) Baa average corporate bond yields and converted into return data using Shiller's (1979) approximation formula.

The second group collects variables that measure macroeconomic risks and that are typical of the literature These factors are used as proxies for systematic, economy-wide risks. In practice, we employ nine factors:[10] the excess return on a value-weighted index portfolio (r_t^M) that includes all stocks traded on the NYSE, AMEX, and Nasdaq (from CRSP); changes in the default risk premium (def_t) measured as the difference between Baa Moody's *yields* and yields on 10-year Treasuries; the change in the term premium ($\Delta term_t$), the difference between 10-year and 1-month riskless yields; the unexpected inflation rate ($UInfl_t$), computed as the residual of an ARMA(1,1) model applied to (seasonally adjusted) CPI inflation; the rate of growth of (seasonally adjusted) industrial production (IP_t); the rate of growth of (seasonally adjusted) real personal consumption growth (PC_t); the 1-month real T-bill return computed as the difference between the 1-month T-bill nominal return and realized CPI inflation (not seasonally adjusted); the traded Liquidity factor (Liq_t) from Pastor and Stambaugh (2003); the bond factor ($Bondp_t$) in Cochrane and Piazzesi (2005).[11] Using a relatively large number of pre-selected factors is typical of the literature.[12] Table 2.1 reports summary statistics.

[10]We do not attempt to either select or "optimize" the number of factors, or to compute linear combinations that may maximize their forecasting power, for instance, as in Zhou (1999) who also adopts one-step methods, but in a GMM framework. The presence of market portfolio returns among the "macro-style" factors may also be justifed by our desire to test whether macroeconomic aggregates represent useful factors that are not already embedded in financial prices.

[11]The traded liquidity factor consists of value-weighted returns on a high-minus-low exposure portfolio on an aggregate liquidity risk factor that sorts stocks on the basis of liquidity risk measures. Næ s and Ø degaard (2011) show the existence of strong linkages between stock market liquidity and business cycle-related macroeconomic aggregates. The bond risk premium factor is constructed as the projection of the equally weighted average of one-year excess holding period return on bonds with maturities of two, three, four, and five years on a constant, the one-year yield, and the two- through five-year forward rates. The bond risk factor is the fitted value of this regression. Cochrane and Piazzesi (2005) and Ludvigson and Ng (2009) investigate the relationship beween this factor and macroeconomic aggregates, following the intuition of Harvey (1989) on the links between the term structure and consumption growth.

[12]For instance Mei (1993) uses five factors; Connor and Korajczyk (1998) find there are more than five factors (although factors in excess of five generally do not play an important role, although they are statistically significant); Ludvigson and Ng (2009) find evidence in favor of eight latent factors. Another strand of the literature has instead used statistical methods to extract latent factors, for instance as principal components, see e.g., Wu and Zhang (2008).

2.4.2 Time-Varying Betas

As an initial way to assess the plausibility of our results, Figure B.1 reports the average (of posterior medians over time) probabilities over our sample of observing a break in the factor loadings, in addition to the intercept, across two different specifications, namely the B-TVB-SV and the homoskedastic B-TVB, for the 23 test assets/portfolios. Clearly the presence of breaks in the idiosyncratic variance process makes a difference in capturing any instability in portfolio betas. Under the B-TVP model the average probability of observing a break is around 40% for the intercept (labeled as factor 1 in the figure) of all portfolios examined, and ranges from 20% for the credit and term spreads (factors 2 and 8) to almost 40% for the bond factor (factor 9). This shows that infrequent and large breaks in betas (as well as Jensen's alphas) can be isolated by the Gibbs sampling algorithm. Under a B-TVB specification, instead, the degree of instability in the factor loadings dramatically collapses. The average probability of a break in betas is around 5% across all risk for the industry portfolios (1-10), while for both the size-sorted and bond portfolios, the average break probability over the sample increases to between 20 and 30% across factors.

Figures B.4-B.5 plot a selection of time series medians and 95% Bayesian credibility intervals computed from the posterior densities of the loadings $\beta_{ij,t}$, obtained from the B-TVB-SV model. To save space, we report plots of time series of risk exposures for all the 23 portfolios used in our estimation, but only for five out of nine specific factors: the U.S. market portfolio, the term spread, industrial production growth, the real T-Bill, and unexpected inflation. Other, similar plots concerning the remaining risk factors—the credit spread, the real consumption growth, the bond and the liquidity factors —are available upon request even though we summarize their contents and implications below. An overview of the plots immediately reveals that the Bayesian estimates of the loadings for all but the market portfolio and the bond risk factor, imply a time path of the factor loadings that is rather smooth over time. This is a first interesting result: even though

(2.5) formally allows factor exposures to be subject to "jumps" over time, as a result of the realization of a latent binary random variable, the resulting posterior densities are actually smoother than what one could retrieve using, say, a naïve rolling window scheme.

For instance, this is evident from a comparison of Figures B.4 and B.3, where in the latter we plot estimated, 5-year rolling window F-MB betas for the 23 test portfolios vs. the market portfolio.[13] Interestingly, this smoothness mimics exactly what many earlier chapters have imposed by assuming near unit root processes ($\beta_{ij,t} = \phi_{ij}\beta_{ij,t-1} + u_{ij,t}$) with small variance of the shocks, but is derived endogenously, as required by the data, which means that occasional large jumps in exposures and/or high volatility of the corresponding process may be accommodated. Second, with a limited number of exceptions that will noted below, the 95% confidence bands are relatively tight, which means that the betas are estimated with a fairly high level of reliability.

In particular, Figure B.4, concerning exposures to market risk, collects most of the loadings for which we have evidence that betas are non-zero. All equity portfolios are characterized by positive and reliably estimated betas. This is not the case for the bond portfolios which essentially show zero exposure to the market risk factor.

As already mentioned, Figure B.3 offers an opportunity to compare the B-TVB-SV estimates with market beta exposures under a the classical F-MB approach described in Appendix A. In Figure B.4, concerning the betas vs. term premium shocks, most equity portfolios are significantly exposed to term premium risk, in the sense that their 95% bands do not systematically include zero. In the figure, the betas fluctuate considerably over the sample period and often change sign. Such betas tend to drift down and to be (significantly) negative for low-decile size-sorted portfolios (i.e., small and medium stocks), for high-tech stocks (especially after 1994), and for junk corporate bonds (at least in the 1980s and 1990s); they are instead positive and often significant in the case

[13] To save space, we do not report all plots of the time-varying, 5-year rolling window betas obtained using a classical two-step estimation scheme, as we describe in Appendix A. These plots are available upon request.

of energy and health stocks.

The plots of time-varying exposures to real output (industrial production growth) risk in Figure B.5 show occasionally larg(er) 95% credibility regions that tend to widen over the sample. However, also in this case, for a large sub-set of portfolios, the corresponding betas are estimated to be negative and significant (nondurables, durables, manufacturing, high-tech, shops, health, and small- and medium-size equity portfolios), while for other portfolios the exposure is positive and significant (energy and utility stocks). Of course, negative exposures to output risk are partially surprising, but because in our model, factors have not been orthogonalized one vs. the others—that will require selecting and imposing a triangular structure that would prove to be "ad hoc"—betas only capture partial effects, after other exposures to business cycle risks are taken into account (see Kramer 1994). An unreported figure concerning betas vs. the short term real rate shows instead exposures that are small and for which the 95% credibility bands tend to include zero for most of the sample. However, close attention reveals that a number of smooth patterns of fluctuations imply sub-periods in which exposures to real rate risk have also been precisely estimated, for instance a negative exposure in the case of durables, high tech, retail shop, first capitalization decile stocks, and medium-term Treasury notes, in the early and mid-1990s.

We have also inspected the remaining four sets of plots concerning the other risk factors (default spread, real consumption growth, liquidity, unexpected inflation and bond factors). With reference to the inflation risk factors, the plots show that even though confidence bands tend to be wide for this factor than for other factors that we have described before, for many portfolios there tends to be still significant evidence of a significantly positive exposure. In contrast, in the case of default spread and real consumption growth risks, all betas imply low variability and narrow 95% credibility regions, but these also fluctuate steadily around zero for the *all* 23 portfolio investigated. The finding is similar in the case of betas vs. the liquidity factor, although these show a stronger temporal variation, which means that sub-samples (sometimes of a length of almost a decade) can be isolated in which this factor commands precisely estimated exposures. When this occurs, betas are mostly positive (for instance, durable stocks in the late 1990s and high-tech stocks in the early 1990s). Finally, although unstable, the results concerning the exposures to Cochrane and Piazzesi's factor are interesting. As one would expect, all bond portfolios show periods in which they have large and significant loadings on this factor, especially during the 1990s and 2001-2004. Equities show more dispersed posteriors which however become large during the 1990s.

Figure B.6 reports posterior medians and 95% credibility intervals for the $\beta_{i0,t}$s estimated from the B-TVB-SV model. In an ICAPM interpretation of (2.1), made possible by the fact that all factors are traded, the time series (of posterior medians for) $\{\beta_{i0,t}\}$ gives indications on mis-pricing. Out of 23 portfolios, in no case the estimated mis-pricing indicators are systematically elevated (in absolute) value. In fact, apart from occasional fluctuations, separate calculations show that the 95% credibility regions include a zero mis-pricing in more than three-quarters of our sample. This is an indication that in its B-TVB-SV implementation, the model (2.4)- (2.5) is well-specified in an economic sense, as it does not imply any evidence of a systematic mis-pricing. Of course, in the case of many portfolios, occasional periods in which the posterior of $\beta_{i0,t}$ fails to include a zero mis-pricing can be found. For instance, there is evidence that all bond portfolios implied positive and tight posteriors for the Jensen's alphas between 2000 and 2004; high-tech and telecommunication stocks were all giving large and significant alphas during the early- to mid-1990s.

Interestingly, Figure B.7 shows that this is not the case when we plot and examine estimates of the $\beta_{i0,t}$ coefficients from a 5-year F-MB implementation. The corresponding Jensen's alphas are strongly fluctuating often reaching extreme levels of $\pm 4\%$ *per month*, i.e., values that are hardly plausible in an asset pricing perspective, and they often appear to be statistically significant in the sense that their 95% confidence bands fail to include a zero mis-pricing. This is of course hardly credible and points more to a failure of the two-pass, rolling window approach than to a misspecification issue with the model in (2.1).

2.4.3 Dynamics in Idiosyncratic Risk

A growing literature (see e.g., Campbell and Xu 2001) has stressed that the idiosyncratic variance of the excess returns on most test portfolios, σ_{it}^2, has undergone important shifts and/or dynamics over the last two decades. We have inspected the filtered GARCH values of residual variance obtained from the classical, two-pass F-MB method (unreported to save space). The presence of rich dynamics is obvious for all the portfolios. In some cases such dynamics turned out to be hard to interpret. First, idiosyncratic variance tends to be large for most portfolios investigated, a sign that the two-pass method provides a very rough explanation to the 23 time series of excess returns. The ratio between the average of the two-pass GARCH estimates $\overline{\hat{\sigma}_{it}^2}$ and the sample variance of portfolio returns over our sample reveals that for a large fraction (13) of the 23 portfolios investigated, this is close to or in excess of 0.5, with all the bond portfolios with ratios in excess of one (because of the rolling window nature of the $\{\hat{\sigma}_{it}\}$, this ratio may actually exceed one). This means that the non-systematic component of excess returns still explains at least 50% of the total variance. Second, most equity portfolios (in practice, all the industry portfolios and size deciles 1-8) record a peak in idiosyncratic variance between 2000 and 2003. In some cases, the rolling window idiosyncratic variance practically doubles between 1999 and 2004, meaning that the MFAPM loses most of its ability to fit excess returns using risk exposures.

Figure B.8 plots posterior medians for σ_{it}^2 estimated from (2.4)- (2.5), along with 95% Bayesian intervals. There are evident spikes in idiosyncratic volatility in the early 2000s and weaker signs of a growing trend towards the end of our sample. The financial crisis of 2008-2009 induces a residual risk increase, but this appears to be minor compared to 1999-

2001, when the model had temporarily lost its ability to fit the U.S. cross-section. In this respect, the fact that the model is more at trouble with the tech stock bust than with the U.S. subprime and credit crunch crises is intriguing. However, the fact that idiosyncratic is countercyclical was largely expected in the light of the literature (see Campbell and Xu 2001). The B-TVB-SV model explains away almost all the variability in excess returns in the case of medium and large cap stocks, and to some extent also government bonds. Spikes in idiosyncratic risk are instead more pronounced for small caps and for a number of industry portfolios, that are explained much less accurately than size-sorted portfolios are. Yet, no clear trend is observed, which is consistent with the more recent evidence reported by Bekaert, Hodrick, and Zhang (2012).

2.4.4 Time-Varying Risk Premia

Table 2.2 reports summary statistics for posterior estimates of the risk premia $\{\hat{\lambda}_{j,t}\}$ $(j = 1, ..., K)$ from the B-TVB-SV model as well as the B-TVP and the homoskedastic B-TVB frameworks. As a benchmark, the table also shows estimates from the second-pass F-MB approach. In the table, we also report the empirical (over the entire 1982-2011 period) standard error for the sample mean of each of the $\lambda_{j,t}$s. From the first panel of the table, it is clear that the classical estimation procedure that non-parametrically tracks time-variation using 5-year rolling windows delivers economically weak implications: only two factors were accurately priced in the cross-section (the market and bond factors), but the former with a p-value exceeding the standard 0.05 threshold and the latter with a difficult to understand, negative sign; moreover, the time series mean of $\hat{\lambda}_{0,t}$ is large, positive (0.29% per month), and statistically significant (its p-value is 0.034), which is problematic to our MFAPM because a non-zero average $\lambda_{0,t}$ implies that omitted risk factors with non-zero risk premia are absorbed by the residual mean.

Therefore the background to our dynamic, state-space results is that a simple, *ad-hoc* rolling window implementation of the MFAPM in (2.4)- (2.5) would yield an embarrassing rejection of the model, in spite of the fact that we are employing as many as nine factors, some of them coming with a strong endorsement of cross-sectional explanatory power from the asset pricing literature. A much more comforting picture emerges from the B-TVB-SV model: the Bayesian design gives evidence of precisely estimated market, liquidity, and macroeconomic (as captured by IP growth shocks) risk premia, with the correct, positive signs (0.339, 0.317, and 0.002 percent per month/unit of risk, respectively). Also the unexpected inflation risk premium is precisely estimated but with a negative sign, similarly to Chen and Ross (1986), Ferson and Harvey (1991), and Lamont (2001). Importantly, the average of the posteriors for $\lambda_{0,t}$ reveals that the intercept is not significantly different from zero.

These results show that while in a F-MB implementation all one gets is evidence that a standard MFAPM is rejected with reference to a typical set of U.S. financial portfolios, such finding is replaced by reassuring evidence that not only the market portfolio (as typical of textbook CAPM) but also a number of macroeconomic factors carry precisely estimated and economically meaningful risk prices. Such empirical findings are less comforting when we impose restrictions on B-TVB-SV. In both cases, the average of the posteriors for $\lambda_{0,t}$ has a mean that is significantly positive, with p-values below 0.05. While in the B-TVP case at least 3 of the 4 factors that commanded positive and significant risk premia in B-TVB-SV set up, in the homoskedastic B-TVB model only market risks appear to be (barely) priced in the U.S. cross-section of stocks and bonds. This is indicative that the restrictions imposed by the B-TVP and the homoskedastic B-TVB models are rejected, an aspect that is investigated in more detail in Section 4.5.

2.4.5 Model Comparisons: Marginal Likelihood Evidence

Similarly to Nardari and Scruggs (2007), we use the marginal likelihood of different models to perform a comparison able to take into account their overall (in-sample) statistical performance, and not only their asset pricing plausibility as in Sections 4.2-4.4. The marginal likelihood of a model is known to take into account both the uncertainty about the size and the presence of structural breaks and the uncertainty concerning the parameters in (2.4)- (2.7). The marginal likelihood of each model is computed as

$$p(\mathbf{R}|\mathbf{F};\mathcal{M}_i) = \int ... \int \sum_{\mathcal{K}} p(\mathbf{R}|\mathbf{B},\mathcal{K},\mathbf{\Sigma},\lambda,\theta,\mathbf{F};\mathcal{M}_i)\times p(\theta,\mathbf{B},\mathcal{K},\mathbf{\Sigma},\lambda|\mathbf{R},\mathbf{F};\mathcal{M}_i)d\mathbf{B}d\mathbf{\Sigma}d\theta d\mathbf{\Sigma}, \tag{2.12}$$

where $\mathcal{M}_i$ identifies the ith model and the posterior density $p(\mathbf{R},\mathbf{B},\mathcal{K},\mathbf{\Sigma},\lambda|\theta,\mathbf{F};\mathcal{M}_i)$ is given by (2.11). Following Chib (1995), we compute the marginal likelihood by replacing the unobservable breaks and parameters in the likelihood of the data generating process defined by (2.4)- (2.7) for each draw.

Table 2.3 reports the marginal (log)likelihoods for each of the model specifications as well as the Bayes factors, the difference between model-specific (log)likelihoods, used as an model selection indicator that naturally penalizes for the different size/complexity of different models (see Kass and Raftery 1995), for each of the alternative frameworks including the two-step F-MB approach, vs. B-TVB-SV. Because Bayes factors are constructed from marginal likelihoods, they measure a model ability to explain the entire distribution (not just first moments) of test asset returns. Bayes factors also permit the simultaneous comparison of multiple models, regardless of whether the models are nested. Because factors and not posterior odds ratios are reported, we implicitly interpret Bayes factors under the assumption that all models are equally likely a priori, as in Nardari and Scruggs (2007). However, to favor interpretations, also the (log)likelihood contributions by each of the 23 test portfolios under each of the models and the corresponding Bayes factors have been computed. Interestingly the B-TVB-SV model shows the higher marginal (log) likelihood values across all of the portfolios under consideration, as well as the higher overall marginal likelihood. By exceeding 100, all the overall Bayes factors are highly significant (see Kass and Raftery (1995), for a justification of this scale/threshold). In particular, the factors vs. the B-TVP and the two-step F-MB implementations are

892 and 5602, respectively, and therefore appear to be decisively in favor of the complete B-TVB-SV framework. The Bayes factor vs. the B-TVB model with stochastic volatility is instead 191 and remains favorable to B-TVB-SV. Surprisingly, the B-TVP model ranks second both in overall terms and for all the test portfolios, thus outperforming the homoskedastic B-TVB alternative. This result emphasizes that by fully acknowledging instability (heteroskedasticity) in idiosyncratic risk (similarly to Nardari and Scruggs 2007) plays a key role beyond that of capturing breaks in the betas. As one would expect, given its ingenious but ad-hoc nature, the classical two-step F-MB approach ranks last with an overall marginal likelihood around 15 times lower than under the B-TVB-SV model. The dominance of the B-TVB-SV framework occurs across all portfolios, but appears to be particularly elevated in the case of bonds and medium and large caps portfolios of stocks.

2.5 Economic Implications

So far our discussion has focussed on the statistical performance of the models in terms of whether there was evidence of either the $\lambda_{0,t}$s or the $\beta_{i0,t}$s coefficients being different from zero and especially with emphasis on the comparison of marginal log-likelihood values. We have concluded that (2.1)-(2.2) is rejected in its standard, two-pass F-MB implementation. However, there was encouraging evidence that the B-TVB-SV model may be less at odds with the data. The results concerning B-TVP have shown that while there are some degrees of freedom as to the way one ought to best model time-variation in risk exposures, capturing instability in stochastic volatility is truly fundamental. Yet, we still know little about the economic implications of the B-TVB-SV framework. In this section, we have assembled some evidence of the economic importance of the estimates of the B-TVB-SV model. In Section 5.1 we tabulate and comment the variance ratios, $VR1$ and $VR2$ described in Appendix C (see also Ferson and Harvey 1991) that measure the degree of misspecification of a MFAPM. Again following Appendix C, Section 5.2 comments on a decomposition of the sources of predictable variation of excess returns due to the MFAPM. Section 5.3 reports pricing tests that measure the quality of the approximation provided by the MFAPM.

2.5.1 Variance Ratios

With reference to the estimates obtained for (2.4)- (2.7), we have computed the (posterior distributions of the) VR1 and VR2 ratios defined in Appendix C and typical of the linear factor model literature. Given their popularity, we just remind a Reader that VR1 should be equal to 1 if the multi-factor model is correctly specified, which means that all the predictable variation in excess returns is captured by variation in risk compensations; at the same time, VR2 should be equal to zero if the multi-factor model is correctly specified. Moreover, as explained in Appendix C, $VR1 = 1$ does not imply that $VR2 = 0$ and viceversa. In what follows, the information at time $t-1$ ($\mathbf{Z}_{t-1}$) used to tease out the total predictable variation in excess returns used as a "denominator" in the empirical results that follow is proxied by the instrumental variables listed in Table 1, plus a dummy variable to account for the so-called "January effect". Note that these instruments fail to simply correspond to the macroeconomic risk factors also because the former do not need to be tradeable portfolios.

Columns 4 and 7 of Table 2.4 present posterior medians of (normalized, by dividing the posterior medians by the variance of the underlying excess return series) $VR1$ and $VR2$ obtained from the B-TVB-SV model for each of the 23 portfolios. Variance ratio results are encouraging. Under a VR1 perspective, we can claim that on average approximately 80% of the predictable variation in excess returns is captured by the B-TVB-SV model. Such a statistic is only 51% in the case of the F-MB implementation (column 1) and goes as low as 47 and 43% for the B-TVP and homoskedastic B-TBV models, respectively. Although in the light of the earlier marginal log-likelihood evidence, this is relatively unsurprising, because the mapping between the ability to capture any predictable variation and the log-likelihood is a complex one, this evidence remains economically important. However even the generally high VR1 ratios from the B-TVB-SV model vary considerably across different test assets. The ratios are relatively high, also in relation to what is typically reported in the literature (see Ferson and Harvey 1991), in the case of government bond portfolios (possibly because we have used Cochrane and Piazzesi's factor) and for a few industries, such as manufacturing, energy, and high-tech, for which VR1 exceeds 90%. It is instead below 50% in the case of the smallest capitalization decile and of non-investment grade corporate bonds, exactly where one would expect our macroeconomic risk factors to have more trouble at fitting the variation in excess returns.[14]

Because $VR1 + VR2 = 1$ does not hold, the finding of good VR1 ratios fails to imply that the VR2 ratios are as close to zero as much as we would want. Yet, VR2 is on average just above 20% in the case B-TVB-SV, to be contrasted with averages across test portfolios of 48-54% in the case of other models/estimation approaches. Moreover, in the case of the B-TVB-SV framework, we record VR2 ratios equal to or inferior to 15% in 9 out of 23 portfolios, which is impressive. All in all, under both the VR1 and VR2, we find evidence of appreciable performance of the model.

2.5.2 Sources of Risk

We have also followed Appendix C and computed the contribution of each factor to the fit offered by the B-TVB-SV model to fitting the predictable variation in excess stock returns. Table 2.5 reports the results.

The highest contribution is given by the market risk factor: with three exceptions (energy, heath, and utility stocks), all the ratios

$$\frac{Var[P(\lambda_{MKT,t}\beta_{iMKT,t|t-1}|\mathbf{Z}_{t-1})]}{Var[P(\sum_{j=1}^{9}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1})]}$$

[14]The 95% credibility regions do little to cast any doubts on this obvious outperformance of the B-TVB-SV framework over the competing models. For instance, the 2.5% posterior lower bound in the case of B-TVB-SV is 62.2% vs. 22.6% in the case of B-TVP and 19.0% in the case of the homoskedastic B-TVB.

concerning stocks exceed 0.5 with peaks in excess of 1 for a number of industries as well as medium-cap portfolios.[15] However, the market factor does not explain most of the predictable variation in excess bond returns, when it is replaced in this leading role by the credit risk factor. As far as stocks are concerned, the next most important contributions come from unexpected inflation (especially for bond and selected industry portfolios) and to some extent, real consumption growth risk, although the heterogeneity across portfolios is pronounced. In the case of bond portfolios, most predictable variation is explained, after taking into credit risk exposures into account, by unexpected inflation, economy-wide market risks, and Cochrane and Piazzesi's specific factor; interestingly, the contribution of yield curve shocks to priced risk is limited. It should not be surprising that some important sources of risk may fail to imply precisely estimated risk premia, as this is a well-known finding in the literature (see e.g., Chan and Lakonishok 1998). Finally, in spite their prices were often precisely estimated, the liquidity and bond risk factors appear to give modest contributions to explain predictable variation in excess stock returns.

2.5.3 Pricing Errors

We follow Geweke and Zhou (1996) and measure the closeness of the pricing approximation provided by an approximate version of (2.8), $E_{t-1}[r_{i,t}] \simeq \lambda_{0,t} + \sum_{j=1}^{K} \beta_{ij,t}\lambda_{j,t}$, by computing at each time t the average squared recursive pricing error across all the N test assets/portfolios,

$$Q_{t,N}^{2} = \frac{1}{N}\left[\beta_{0,t}'\left(\mathbf{I}_N - \mathbf{B}_t\left(\mathbf{B}_t'\mathbf{B}_t\right)^{-1}\mathbf{B}_t'\right)\beta_{0,t}\right] \qquad t = 1, ..., T, \tag{2.13}$$

where $\beta_{0,t}$ is the $N \times 1$ vector of intercepts, $\mathbf{I}_N$ is an N-dimensional identity matrix, and $\mathbf{B}_t \equiv (\iota_N, \beta_{1,t}, ..., \beta_{K,t})$ is a $N \times K$ matrix collecting vectors of time t betas of all the assets/portfolios vs. each of the K risk factors, with $\beta_{j,t} \equiv (\beta_{1j,t}, ..., \beta_{Nj,t})'$ a $N \times 1$ vector

[15] These ratios may exceed 100% because $Var[P(\sum_{j=1}^{9} \lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1})]$ will also reflect the contribution of covariance terms between factor terms.

of factor loadings on the jth risk factor. These pricing errors are recursive because at each point in time they are obtained using only information available up to that point. Because our Gibbs sampling scheme allows to derive posteriors for all the objects that enter $\beta_{0,t}$ and $\mathbf{B}_t$, we also compute the posterior density of the average (squared) pricing error statistic.

Table 2.6 reports the average monthly pricing errors, $Q_{t,N}$, for each of our models across different sub-samples. With reference to the full-sample, the B-TVB-SV model yields both the lowest average pricing error (0.21% per month) and the lowest median posterior error (0.19%). Such statistics are practically between one-half and two-thirds those that one would obtain under a B-TVB homoskedastic model (0.41 and 0.35 percent, respectively). Interestingly, the B-TVP model seems to fit the data well on the basis of the Bayes odds ratio in Table 2.3, but fails to price our test portfolios (it gives average and median posterior errors of 0.54 and 0.51 percent, respectively) as accurately as the homoskedastic B-TVB model does. The performance of the classical two-step F-MB scheme is poor, yielding average and median pricing errors of 0.63%. Moreover, B-TVB-SV consistently outperforms all other models in all sub-samples. Its advantage is always substantial in the sense that B-TVB-SV always cuts the average error of the second best model by at least 40%. Interestingly, the pricing errors tend to increase over our sample, especially when one compares the 1982-1998 with the 1999-2011 interval. However, there is no evidence of errors during the recent financial crisis being systematically higher than in the overall 1999-2011 sample.

Figure B.9 plots the time series of average pricing errors $Q_{t,N}$ for all the models (top panel) and only for the B-TVB-SV and the B-TVP cases, rescaling the errors from the former model (on the right axis) to better emphasize similarities and differences (bottom panel). The top panel shows that, apart from a short period in early 1993 (when B-TVB became competitive), B-TVB-SV gave uniformly lower average pricing errors than all other models. The F-MB scheme gives uniformly high but constant average errors. The

B-TVP model gives a highly variable performance, with enormous spikes of mis-pricing around 1993, in 1999-2000, and during the financial crisis. The bottom panel of the figure shows that the dynamics of pricing errors under B-TVB-SV and B-TVP—both models including a stochastic volatility component—are not that dissimilar, in the sense that also errors from B-TVB-SV spike up in 1993, 1999-2001 and during 2011. However, the more parsimonious dynamics imposed by infrequent, large structural breaks under B-TVB-SV reduces the pricing errors also keeping the latter more stable across our sample.

2.6 Discussion and Conclusion

We conclude with a few remarks on the qualitative differences when estimation has been repeated after changing the choices that have driven analysis so far. In general we have found modest differences and this provides further support to the encouraging results revealed by Sections 4 and 5. To save space, we have not plotted or tabulated complete set of results, that remain available from the Authors.

As explained in Section 2, our Bayesian implementation tries to avoid the generated regressor problems that have plagued F-MB's methods. In practice, this means that we have considered the full joint posterior distribution of the time $t-1$ (expected) factor sensitivities $\beta_{ij,t|t-1}$ and the corresponding risk premia $\lambda_{j,t}$ when estimating equilibrium asset pricing restrictions. In spite of its internal consistency and the fact this appears to be the only way in which all relevant (estimation) uncertainty may be taken into account, to see whether our economic insights may be mostly driven by this approach, we have repeated the entire analysis under the B-BTV-SV model by using the simpler and traditional method of using plain (median, in our case) betas from the first pass rather than their full posterior distribution at time t, as in Jostova and Philipov (2005). This "mixed" B-TVB-SV/F-MB implementation reveals substantial over-pricing of risks, in the sense that the average $\lambda_{0,t}$ is precisely estimated (the 95% confidence bands for $\lambda_{0,t}$ stop including zero), so that this hybrid framework would be rejected. This result establishes that—even assuming an identical B-BTV-SV—it is its consistent implementation that takes into account all statistical uncertainty that is key.

Additionally, we have experimented with an informative prior in the second pass in order to put some structure (constraints) on the distribution and moments of the risk premia. These are now postulated to be normally distributed with zero mean and variance such that there is 95% probability that annualized premia are smaller in absolute value than the larger between the absolute value of the minimum and the maximum return observed in the sample for all the assets. We record a striking reduction in the variability of the estimated posterior distributions (as well as their medians) for the risk premia relative to the baseline case. The qualitative results and insights from Table 2 in the B-BTV-SV case apply intact and, in general, using informative priors on the premia constrains their variability, so we find both less variable premia: industrial production growth, liquidity, inflation, and especially market risks are important drivers of the cross section of U.S. stock and bond returns.

We have proposed a new way to parameterize and estimate in state-space form a typical macro-based MFAPM with time-varying risk exposures and premia. Our Bayesian approach is based on a formal modelling of the latent process followed by risk exposures and idiosyncratic volatility capable to capture structural shifts in parameters. This method can also be interpreted as a novel way to overcome the two-pass approach advocated by Fama and MacBeth (1973) and used in a substantive body of applied work in finance. Given a general B-TVB-SV framework, we have also considered special cases that are obtained by imposing restrictions, and in particular a B-TVP model in which betas change continuously but in small amounts, and a homoskedastic B-TVB model in which volatility is constant.

Our application to monthly, 1972-2011 U.S. stock and bond returns shows that the empirical implications of a Bayesian state-space implementation of (2.4)- (2.5) are plausible and there are indications that the model is consistent with the data. For instance, most portfolios do not appear to have been grossly mis-priced and a few risk premia are precisely estimated with a plausible sign. Market, liquidity, and industrial production (real output) growth risk are significantly priced. Bayes odds ratios and marginal likelihood comparisons indicate that the B-TVB-SV outperforms both the two-step F-MB and the homoskedastic B-TVB models. The heteroskedastic B-TVP appears to be closer to the full-scale B-TVB-SV one. This confirms the findings in Nardari and Scruggs (2007), here extended to the case of time-varying betas. However, an analysis of the average pricing errors shows that large but infrequent breaks in factor exposures are considerably more successful. Finally, the finding that the heteroskedastic B-TVB models ranks second below B-TVB-SV is a powerful indication of the importance to explicitly model stochastic volatility when implementing macro-driven multi-factor asset pricing models.

Appendices

Appendix B

B.1 Traditional Two-Stage Fama-MacBeth Estimation

Here we outline the classical, two-stage procedure *à* la Fama and MacBeth (1973) also used by Ferson and Harvey (1991) and very popular in the empirical finance literature. This represents the standard benchmark for the estimation of the equilibrium asset pricing model shown in (1)-(2) in the chapter.

In the first stage, for each of the assets, the factor betas are estimated using time-series regressions from historical excess returns on the assets and economic factors. That is, for month t, (1) in the main text is estimated using the previous sixty months (ranging from $t-61$ to $t-1$) in order to obtain estimates for the betas, $\hat{\beta}^{60}_{ij,t}$. This time-series regression is updated each month. The choice of a 60-month rolling window scheme is typical of the literature.[1] In the second stage, the equilibrium restriction (2) in the main text is estimated for each of the periods in our sample a cross-sectional regression using ex-post

[1] See Chen, Roll and Ross (1986) and Ferson and Harvey (1991). In unreported tests we have attempted to optimize this choice by picking the sliding window that produced the lowest average information criterion, such as the BIC. We find that a 5-year window gives at all times a BIC which is sensible lower than any other window in the range [3, 10] years.

realized excess returns:

$$r_{i,t} = \lambda_{0,t} + \sum_{j=1}^{K} \lambda_{j,t}\hat{\beta}_{ij,t}^{60} + \zeta_{i,t} \qquad i = 1, ..., N, \quad t = 61, ..., T. \tag{B.1}$$

Clearly, this T cross-sectional regressions simply implement (2) in the main text in a nonparametric fashion, in the sense that any resulting time variation in the $\lambda_{0,t}$ and $\lambda_{j,t}$ coefficients fails to be explicitly and parametrically related to any of the instruments assumed by the researcher, even though additional projections/regressions remain possible. In (18) in the main text $\lambda_{0,t}$ is the zero-beta (abnormal) excess return and the $\lambda_{j,t}$s are proxies for the factor risk premiums on each month, $j = 1, ..., K$. This derives from the fact that if one considers a portfolio κ such that $\hat{\beta}_{\kappa j,t}^{60} = 0$ for all $j \neq \kappa$ and $\hat{\beta}_{\kappa\kappa,t}^{60} = 1$, then $\lambda_{\kappa,t}$ is simply the conditional mean of $r_{\kappa,t} - \lambda_{0,t}$. Notice that $\lambda_{0,t}$ should equal zero $\forall t$ if the model is correctly specified, because in the absence of arbitrage all zero-beta assets should command a rate of return that equals the short-term rate. Tests of multi-factor models evaluate the importance of the economic risk variables by evaluating whether their risk premiums are priced or whether, on average, the (second-stage, estimated) coefficients $\hat{\lambda}_{j,t}$ are significantly different from zero.

B.2 The Gibbs Sampling Algorithm

In this section we derive the full conditional posterior distributions of the latent variables and the model parameters discussed in Section 2.4. For the ease of exposition we report the results for the ith asset. We use a burn-in period of 1,000 and draw 5,000 observations storing every other of them to simulate the posterior distribution of parameters and latent variables. The resulting autocorrelations of the draws are very low.[2]

[2] In order to gain a rough idea of how well the chain mixes in our algorithm we follow Primiceri (2005) and check the autocorrelation function of the draws.

B.2.1 Step 1. Sampling K_β

The structural breaks in the conditional dynamics of the factor loadings measured by the latent binary state $\kappa_{i0t}, ..., \kappa_{iKt}$, are drawn using the algorithm of Gerlach et al. (2000). This algorithm increases the efficiency of the sampling procedure since allows to generate $\kappa_{it} = (\kappa_{i0t,...,}\kappa_{iKt})$, without conditioning on the relative regression parameters $\beta_{it} = (\beta_{i0t}, ..., \beta_{iKt})$. The conditional posterior density of κ_{it}, $t = 1, ..., T, i = 1, ..., N$, for each of ith asset/portfolio is defined as

$$
\begin{aligned}
& p\left(\kappa_{i0t,...,}\kappa_{iKt}|\mathcal{K}_{i\beta[-t]}, \mathcal{K}_{i\sigma}, \mathbf{\Sigma}_i, \theta_i, \mathbf{R}_i, \mathbf{F}\right) \propto \\
& \propto p\left(\mathbf{R}_i|\mathcal{K}_{i\beta}, \mathcal{K}_{i\sigma}, \mathbf{\Sigma}_i, \theta_i, \mathbf{F}\right) p\left(\kappa_{i0t,...,}\kappa_{iKt}|\mathcal{K}_{i\beta[-t]}, \mathcal{K}_{i\sigma}, \mathbf{\Sigma}_i, \theta_i, \mathbf{F}\right) \\
& \propto p\left(r_{i,t+1,...,}r_{i,T}|r_{i,1}, ..., r_{i,t}, \mathcal{K}_{i\beta}, \mathcal{K}_{i\sigma}, \mathbf{\Sigma}_i, \theta_i, \mathbf{F}\right) p\left(r_{i,t}|r_{i,1,...,}r_{i,t-1}, \mathcal{K}_{i\beta[1:t]}, \mathcal{K}_{i\sigma}, \mathbf{\Sigma}_i, \theta_i, \mathbf{F}\right) \\
& p\left(\kappa_{i0t,...,}\kappa_{iKt}|\mathcal{K}_{i\beta[-t]}, \mathcal{K}_{i\sigma}, \mathbf{\Sigma}_i, \theta_i, \mathbf{F}\right) \qquad \text{(B.2)}
\end{aligned}
$$

where $\mathcal{K}_{i\beta[-t]} = \left\{\{\kappa_{ijs}\}_{j=0}^{K}\right\}_{s=1,s\neq t}^{T}$, $\mathcal{K}_{i\beta[1:t]} = \left\{\{\kappa_{ijd}\}_{j=0}^{K}\right\}_{d=1}^{t}$ and $\mathcal{K}_{i\sigma} = \{\kappa_{i\nu,t}\}_{t=1}^{T}$. We assume that each of the κ_{ijs} breaks are independent from each other such that the joint density is defined as $\prod_{j=0}^{K} \pi_{ij}^{\kappa_{ijs}} (1 - \pi_{ij})^{1-\kappa_{ijs}}$. The remaining densities

$$
p\left(r_{i,t+1,...,}r_{i,T}|\ r_{i,1,...,}r_{i,t},\ \mathcal{K}_{i\beta}, \mathcal{K}_{i\sigma}, \mathbf{\Sigma}_i,\ \theta_i, \mathbf{F}\right)
$$

and

$$
p\left(r_{it}|r_{i1,...,}r_{it-1}, \mathcal{K}_{i\beta t}, \mathcal{K}_{i\sigma}, \mathbf{\Sigma}_i, \theta_i, \mathbf{F}\right)
$$

are evaluated as in Gerlach et al. (2000). Notice that, since κ_{ijt} is a binary state the integrating constant is easily evaluated.

B.2.2 Step 2. Sampling the Factor Loadings B

The full conditional posterior density for the time-varying factor loadings is computed using a standard forward filtering backward sampling as in Carter and Kohn (1994). For each of the $i = 1, ..., N$ assets, the prior distribution of the $\beta_{i0}, ..., \beta_{iK}$ loadings is a multivariate normal with the location parameters corresponding to the OLS parameter estimates and a covariance structure which is diagonal and defined by the variances of the OLS estimates. The initial prior are sequentially updated via the Kalman Filtering recursion, then the parameters are drawn from the posterior distribution which is generated by a standard backward recursion (see Fruhwirth-Schnatter 1994, Carter and Kohn 1994, and West and Harrison 1997).

B.2.3 Step 3 and 4. Sampling the Breaks and the Values of the Idiosyncratic Volatility.

In order to draw the structural breaks $\mathcal{K}_{i\sigma}$ and the idiosyncratic volatilities $\mathbf{\Sigma}_i$ for each of the *ith* portfolios, we follow a similar approach as in step 1. The stochastic breaks $\mathcal{K}_{i\sigma}$ are drawn by using the Gerlach et al. (2000) algorithm. The conditional variances $\ln \sigma_{it}^2$, does not show a linear structure even though still preserving the standard properties of state space models. The model is rewritten as

$$\ln \left(r_{i,t} - \beta_{i0t} - \sum_{j=1}^{K} \beta_{ijt} F_{jt} \right)^2 = \ln \sigma_{it}^2 + u_t$$

$$\ln \sigma_{it}^2 = \ln \sigma_{it-1}^2 + \kappa_{\nu it} \nu_{it} \tag{B.3}$$

where $u_t = \ln \varepsilon_t^2$ has a $\ln \chi^2(1)$. Here we follow Omori et al. (2010) and approximate the $\ln \chi^2(1)$ distribution with a finite mixture of ten normal distributions, such that the density of u_t is given by

$$p(u_t) = \sum_{l=1}^{10} \varphi_l \frac{1}{\sqrt{\varpi_l^2 2\pi}} \exp\left(-\frac{(u_t - \mu_l)^2}{2\varpi_l}\right) \tag{B.4}$$

with $\sum_{l=1}^{10} \varphi_l = 1$. The appropriate values for μ_l, φ_l and ϖ_l^2 can be found in Omori et al. (2010). Mechanically in each step of the Gibbs Samplers we simulate at each time t a component of the mixture. Now, given the mixture component we can apply the standard Kalman filter method, such that $\mathcal{K}_{i\sigma}$ and $\mathbf{\Sigma}_i$ can be sampled in a similar way as $\mathcal{K}_{i\beta[t]}$ and $\beta_{i0[t]}, ..., \beta_{iK[t]}$ in the first and second step. The initial prior of the log idiosyncratic volatility $\ln \sigma_{i0}^2$ is normal with mean -1 and conditional variance equal to 0.1.

B.2.4 Step 5a. Sampling the Time-Varying Risk Premia

The cross-sectional restriction in (2) in the main text is satisfied at each time t conditional on the latent states $\mathbf{B}_{t|t-1} = \left\{\left\{\beta_{ijt|t-1}\right\}_{i=1}^{N}\right\}_{j=0}^{K}$ and $\mathbf{\Sigma}_t = \{\sigma_{it}^2\}_{i=1}^N$. Given an initial normal-inverse gamma prior, the full conditional of the equilibrium risk premia $\lambda_t = (\lambda_{0t}, ..., \lambda_{Kt})$ at time t, is defined as

$$p(\lambda_t|\tau, \mathbf{B}_{t|t-1}, \mathbf{\Sigma}_t, R_t) \propto |\Sigma_t^*|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(R_t - \mu_t^*)^\top (\Sigma_t^*)^{-1} (R_t - \mu_t^*)\right\} \tag{B.5}$$

where $R_t = (r_{1t,....,}r_{N,t})$ and Σ_0, μ_0 respectively the prior mean and variance of λ_t, such that the conditional (ex-ante time-varying) risk premia can be sampled at each time t by a normal distribution with $\mu_t^* = \Sigma_t^*(\Sigma_0^{-1}\mu_0 + \tau^{-2}X_{t-1}^\top R_t)$ and $\Sigma_t^* = (\Sigma_0^{-1} + \tau^{-2}X_{t-1}^\top X_{t-1})^{-1}$, $X_{t-1} = \left[\iota, \mathbf{B}_{t|t-1}\right]$, respectively as location and scale parameters. The conditional posterior for the variance of the risk premia τ^2 is an inverse gamma distribution

$$p(\tau^2|\lambda_t, \mathbf{B}_{t|t-1}, \mathbf{\Sigma}_t, R_t) \propto \tau^{-a_0} \exp\left(-\frac{b_0}{2\tau}\right) \prod_{i=1}^{N} \frac{1}{\tau} \exp\left(\frac{-\left(r_{it} - \lambda_{0t} - \sum_{j=1}^{K} \beta_{ijt|t-1}\lambda_{jt}\right)^2}{2\tau^2}\right) \tag{B.6}$$

such that τ^2 can be sampled from an inverse-gamma distribution with scale parameter $b = b_0 + \sum_{i=1}^{N}(r_{it} - \lambda_{0t} - \sum_{j=1}^{K}\beta_{ijt|t-1}\lambda_{jt})^2$ and degrees of freedom $a = a_0 + N$.

B.2.5 Step 5b. Sampling the Stochastic Breaks Probabilities

The full conditional posterior densities for the breaks probabilities $\pi = (\pi_{i1,...,}\pi_{iK})$ is given by

$$p\left(\pi|q^2, \mathbf{B}, \mathbf{\Sigma}, \mathcal{K}_\beta, \mathbf{R}, \mathbf{F}\right) \propto \prod_{j=0}^{K} \pi_{ij}^{a_{ij}-1}\left(1-\pi_{ij}\right)^{b_{ij}-1} \prod_{t=1}^{T} \pi_{ij}^{\kappa_{ijt}}(1-\pi_{ij})^{1-\kappa_{ijt}} \tag{B.7}$$

and hence the individual π_{ij} parameter can be sampled from a Beta distribution with shape parameters $a_{ij} + \sum_{t=1}^{T}\kappa_{ijt}$ and $b_{ij} + \sum_{t=1}^{T}(1-\kappa_{ijt})$ for $j = 0, ..., K$. Likewise the full conditional posterior distribution for the breaks probabilities in the idiosyncratic volatilities π_ν is given by

$$p\left(\pi_\nu|q^2, \mathbf{B}, \mathbf{\Sigma}, \mathcal{K}_\sigma, \mathbf{R}, \mathbf{F}\right) \propto \pi_{i\nu}^{a_{i\nu}-1}\left(1-\pi_{i\nu}\right)^{b_{i\nu}-1} \prod_{t=1}^{T} \pi_{i\nu}^{\kappa_{i\nu t}}(1-\pi_{i\nu})^{1-\kappa_{i\nu t}}$$

such that the individual $\pi_{i\nu}$ can be sampled from a Beta distribution with shape parameters $a_{i\nu} + \sum_{t=1}^{T}\kappa_{i\nu t}$ and $b_{i\nu} + \sum_{t=1}^{T}(1-\kappa_{i\nu t})$ for $i = 1, ..., N$.

B.2.6 Step 5c. Sampling the Conditional Variance of the States

The prior distributions for the conditional volatilities of the factor loadings β_{ijt} for $j = 0, ..., K$ are inverse-gamma

$$p\left(q_{ij}^2|\pi, \mathbf{B}, \mathbf{\Sigma}, \mathcal{K}_\beta, \mathcal{K}_\sigma, \mathbf{R}, \mathbf{F}\right) \propto q_{ij}^{-\nu_{ij}} \exp\left(-\frac{\delta_{ij}}{2q_{ij}^2}\right) \prod_{t=1}^{T}\left(\frac{1}{q_{ij}}\exp\left(\frac{-\left(\beta_{ijt} - \beta_{ijt-1}\right)^2}{2q_{ij}^2}\right)\right)^{\kappa_{ijt}} \tag{B.8}$$

hence q_{ij}^2 is sampled from an inverse-gamma distribution with scale parameter $\nu_{ij} + \sum_{t=1}^{T} \kappa_{ijt} \left(\beta_{ijt-}\beta_{ijt-1}\right)^2$ and degrees of freedom equal to $\nu_{ij} + \sum_{t=1}^{T} \kappa_{ijt}$. Likewise the full conditional of the variance for the idiosyncratic log volatility $q_{i\nu}^2$ is defined as

$$p\left(q_{i\nu}^2 | \pi, \mathbf{B}, \mathbf{\Sigma}, \mathcal{K}_\beta, \mathcal{K}_\sigma, \mathbf{R}, \mathbf{F}\right) \propto q_{i\nu}^{-\nu_{i\nu}} \exp\left(-\frac{\delta_{i\nu}}{2q_{i\nu}^2}\right) \prod_{t=1}^{T} \left(\frac{1}{q_{i\nu}} \exp\left(\frac{-\left(\ln \sigma_{it}^2 - \ln \sigma_{it-1}^2\right)^2}{2q_{i\nu}^2}\right)\right)^{\kappa_{i\nu t}} \tag{B.9}$$

such that $q_{i\nu}^2$ is sampled from an inverted Gamma distribution with scale parameter $\nu_{i\nu} + \sum_{t=1}^{T} \kappa_{i\nu t}(\ln \sigma_{it}^2 - \ln \sigma_{it-1}^2)^2$ and degrees of freedom equal to $\nu_{i\nu} + \sum_{t=1}^{T} \kappa_{i\nu t}$.

B.3 Choice of Priors

Realistic values for the different prior distributions obviously depend on the problem at hand.[3] In general, we use weak priors, excluding the size of the breaks $\mathbf{Q}_i$ and the probabilities $\Pr(\kappa_{1ij,\tau} = 1)$ and $\Pr(\kappa_{2i,\tau} = 1)$ for which our priors are quite informative. This is also important because these priors restrict the maximum number of breaks of maximum magnitude and therefore help to identify the factor exposures, which is otherwise rather problematic because linear multifactor models are subject to well-known indeterminacy problems upon rotations of factors and risk premia (see e.g., McCulloch and Rossi, 1991). The prior shape parameters for the probability of breaks in the dynamics of the price sensitivities is set to be $a_{ij} = 3.2$ and $b_{ij} = 60$. As such,

$$E\left[\pi_{ij}\right] = \frac{3.2}{3.2 + 60} = 0.05 \qquad \text{and} \qquad Std\left[\pi_{ij}\right] = \left(\frac{3.2 \times 60}{(3.2 + 60)^2(3.2 + 60 + 1)}\right)^{1/2} = 0.03$$

which means an expected 5% prior probability of a random shock in the dynamics of factor loadings. With respect to the idiosyncratic volatility, the shape hyperparameters

[3] Groen, Paap and Ravazzolo (2012) discuss prior sensitivity analysis and MCMC convergence tests, see Appendices C and D, which are also used in this chapter. Specific prior values and results for convergence tests are available upon request.

are set to be $a_{i\nu} = 1$ and $b_{i\nu} = 99$, such that

$$E\left[\pi_{i\nu}\right] = \frac{1}{1+99} = 0.01 \qquad \text{and} \qquad Std\left[\pi_{i\nu}\right] = \left(\frac{99}{100^2 \times 101}\right)^{1/2} = 0.01$$

which set the expected prior probability of having a break in the dynamics of idiosyncratic risks to be equal to 1%. These small prior probabilities makes the modelling dynamics more parsimonious, mitigating the magnitude of prior information, letting the data speak about the likelihood of random breaks. The marginal (expected) posterior probability of random breaks both in the factor loadings and in the alphas are reported in Table B1. The prior beliefs on the size of the breaks are inverse-gamma distributed. The prior scale hyper-parameters $\gamma_{ij}, \gamma_{i\nu}$ and the $\delta_{ij}, \delta_{i\nu}$ degrees of freedom are calibrated supporting a prior view for premiums to be normally distributed with zero mean and variance such that there is 95% probability that annualized premia are smaller in absolute value than the larger between the absolute value between the minimum and the maximum return observed in the sample for all the assets. Finally, the prior residual variance is centered at about 10, a value that appeared in the higher range of the maximum likelihood estimates. All other priors imply that the posteriors tend to be centered around their maximum likelihood estimates which eases comparisons with the existing literature.

B.4 Variance Ratio and Decomposition Tests

We use the posterior densities of the time series of factor loadings and risk premia to perform a number of tests that allow us to assess whether a posited asset pricing framework may explain an adequate percentage of excess asset returns. (9) in the main text decomposes excess asset returns in a component related to risk, represented by the term $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$ plus a residual $\lambda_{0,t} + e_{i,t}$. In principle, a multi-factor model is as good as the implied percentage of total variation in excess returns explained by its first component, $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$. However, here we should recall that even though (9) refers to excess returns, it remains a statistical implementation of the framework in (1) in the chapter. This implies that in practice it may be naive to expect that $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$ be able to explain much of the variability in excess returns. A more sensible goal seems to be that $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$ ought to at least explain the *predictable* variation in excess returns. We therefore follow earlier literature, such as Karolyi and Sanders (1998), and adopt the following approach. First, the excess return on each asset is regressed onto a set of M instrumental variables that proxy for available information at time $t-1$, $\mathbf{Z}_{t-1}$,

$$x_{i,t} = \delta_{i0} + \sum_{m=1}^{M} \delta_{im} Z_{m,t-1} + \xi_{i,t}, \tag{i}$$

to compute the sample variance of fitted values,

$$Var[P(x_{it}|\mathbf{Z}_{t-1})] \equiv Var\left[\widehat{\delta}_{i0} + \sum_{m=1}^{M} \widehat{\delta}_{im} Z_{m,t-1}\right], \tag{B.10}$$

where the notation $P(x_{it}|\mathbf{Z}_{t-1})$ means "linear projection" of x_{it} on a set of instruments, $\mathbf{Z}_{t-1}$. Second, for each asset $i = 1, ..., N$, a time series of fitted (posterior) risk compensations, $\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}$, is regressed onto the instrumental variables,

$$\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1} = \widehat{\delta}'_{i0} + \sum_{m=1}^{M} \widehat{\delta}'_{im} Z_{m,t-1} + \xi'_{i,t} \tag{B.11}$$

to compute the sample variance of fitted risk compensations:

$$Var\left[P\left(\sum_{j=1}^{K} \lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right] \equiv Var\left[\widehat{\delta}'_{i0} + \sum_{m=1}^{M} \widehat{\delta}'_{im} Z_{m,t-1}\right]. \tag{B.12}$$

At this point, it is informative to compute and report two variance ratios, commonly called $VR1$ and $VR2$, after Ferson and Harvey (1991):

$$VR1 \equiv \frac{Var\left[P\left(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]}{Var[P(x_{it}|\mathbf{Z}_{t-1})]} > 0 \quad \text{(p)}$$

$$VR2 \equiv \frac{Var\left[P\left(x_{i,t}-\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]}{Var[P(x_{it}|\mathbf{Z}_{t-1})]} > 0. \quad \text{(B.13)}$$

VR1 should be equal to 1 if the multi-factor model is correctly specified, which means that all the predictable variation in excess returns is captured by variation in risk compensations; at the same time, VR2 should be equal to zero if the multi-factor model is correctly specified. Importantly, when these decomposition tests are implemented using the estimation outputs obtained from our B-TVB-SV framework, drawing from the joint posterior densities of the factor loadings $\beta_{ij,t|t-1}$ and the implied risk premia $\lambda_{j,t}$, $i=1,...,N$, $j=1,...,K$, and $t=1,...,T$, and holding the instruments fixed over time, it is possible to compute VR1 and VR2 in correspondence to each of such draws and hence obtain their posterior distributions.[4]

Finally, the predictable variation of returns due to the multi-factor model may be further decomposed into the components imputed to each of the individual systematic risk factors, by computing the factoring of $Var[P(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1})]$ as

$$\sum_{j=1}^{K} Var\left[P\left(\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right] + \sum_{j=1}^{K}\sum_{k=1}^{K} Cov[P\left(\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right), P\left(\lambda_{k,t}\beta_{ik,t|t-1}|\mathbf{Z}_{t-1}\right)] \quad \text{(B.14)}$$

and tabulating $Var\left[P\left(\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]$ for $j=1,...,K$ as well as the residual factor $\sum_{j=1}^{K}\sum_{k=1}^{K} Cov[\,P\left(\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right), P\left(\lambda_{k,t}\beta_{ik,t|t-1}|\mathbf{Z}_{t-1}\right)]$ to pick up any interaction

[4]Notice that $VR1=1$ does not imply that $VR2=0$ and viceversa, because

$$Var[P(x_{it}|\mathbf{Z}_{t-1})] \neq Var\left[P\left(\sum_{j=1}^{K}\hat{\lambda}_{j,t}\hat{\beta}_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right] + Var\left[P\left(x_{i,t}-\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right].$$

terms. Note that because of the existence of the latter term, the equality

$$\sum_{j=1}^{K} \frac{Var\left[P\left(\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]}{Var\left[P\left(\sum_{j=1}^{K}\lambda_{j,t}\beta_{ij,t|t-1}|\mathbf{Z}_{t-1}\right)\right]} = 1 \tag{B.15}$$

fails to hold, i.e., the sum of the K risk compensations should not equal the total predictable variation from the asset pricing model because of the covariance among individual risk compensations. This derives from the fact that even though in (1) in the chapter the risk factors are assumed to be orthogonal, this does not imply that their time-varying total risk compensations ($\lambda_{j,t}\beta_{ij,t|t-1}$ for $j = 1, ..., K$) should be orthogonal.

B.5 Tables and Figures

Table B.1: Descriptive Statistics

This table reports the descriptive statistics for each of the 23 portfolios used in the empirical analysis as well as the risk factors and the instrumental variables. Data are monthly and cover the sample period 1972:01 - 2011:12.

Portfolio/Factor	Mean	Median	Std. Dev.	Sharpe Ratio
	10 Industry Portfolios, Value-Weighted			
Non-Durable Goods	1.107	1.135	4.473	0.248
Durable Goods	0.809	0.815	6.649	0.122
Manufacturing	0.988	1.195	5.195	0.190
Energy	1.163	0.990	5.672	0.205
High-Tech	0.924	0.950	6.897	0.134
Telecommunications	0.948	1.175	4.891	0.194
Shops and Retail	0.974	1.060	5.447	0.179
Healthcare	0.990	1.050	5.094	0.194
Utilities	0.933	0.995	4.142	0.225
Other	0.871	1.320	5.439	0.160
	10 Size-Sorted Portfolios, Value-Weighted			
Decile 1	1.073	1.205	6.347	0.169
Decile 2	1.083	1.390	6.491	0.167
Decile 3	1.125	1.545	6.162	0.182
Decile 4	1.089	1.500	5.952	0.183
Decile 5	1.127	1.680	5.811	0.194
Decile 6	1.081	1.180	5.412	0.200
Decile 7	1.088	1.255	5.382	0.202
Decile 8	1.024	1.275	5.262	0.195
Decile 9	0.986	1.335	4.853	0.203
Decile 10	0.844	1.075	4.473	0.189
	Bond Returns			
10-Year T-Note	0.679	0.628	2.299	0.295
5-Year T-Note	0.635	0.585	1.629	0.390
Baa Corp. Bond (10-20 years)	0.831	0.863	3.237	0.257
	Economic Risk Factors			
Excess Value-Weighted Mkt	0.452	0.800	4.681	0.097
Default Premium	0.192	0.461	3.481	
Term Spread	0.000	0.000	0.406	
Industrial Prod. Growth	0.186	0.256	0.755	
Real Per-capita Cons. Growth	0.255	0.262	0.338	
Real T-Bill Interest Rate	0.087	0.102	0.357	
Unexpected Inflation	0.000	-0.016	0.301	
Bond Risk Factor	1.093	0.982	1.944	0.562
Liquidity Factor	0.497	0.232	3.621	0.137
	Instrumental Variables			
Term Yield Spread	1.715	1.910	1.329	
Credit Yield Spread	1.111	0.960	0.488	
Dividend Yield	3.029	2.952	1.259	

Table B.2: Risk Premia

This table reports statistics describing the posterior distribution of the risk premia on each factor across different model specifications. Data are monthly and cover the sample period 1972:01 - 2011:12. The first ten years of monthly data are used to calibrate the priors for all the models except for the standard two-step Fama-MacBeth procedure.

	Average	Std. Error	t-stat	p-value	2.5%	50%	97.5%
	Two-step Fama-MacBeth approach						
Intercept	0.2909	0.1363	2.1350	**0.0336**	-3.3346	0.3281	3.3471
Market	0.2593	0.1408	1.8414	**0.0665**	-8.7553	0.6739	7.9319
Credit Spread	0.2208	0.2706	0.8161	0.4151	-4.5672	0.3022	4.8480
Term spread	0.0042	0.0347	0.1201	0.9045	-1.0198	-0.0018	1.1440
IP Growth	-0.0130	0.0092	-1.4086	0.1600	-0.3368	-0.0210	0.3218
Real Consuption Growth	0.0061	0.0039	1.5485	0.1226	-0.1917	-0.0006	0.2309
Real T-bill Rate	-0.0264	0.0412	-0.6414	0.5217	-1.4068	0.0145	1.4703
Unexpected Inflation	-0.0085	0.0062	-1.3670	0.1727	-0.2079	-0.0134	0.2100
Bond Risk Factor	-0.4633	0.1883	-2.4598	**0.0145**	-6.9685	-0.3831	5.1747
Liquidity Factor	0.4012	0.3471	1.1558	0.2487	-12.4488	0.1301	12.9321
	Bayesian model with time-varying betas and idiosyncratic risk						
Intercept	0.4125	0.2924	1.4108	0.1593	0.3406	0.4913	0.6432
Market	0.3391	0.1298	2.6119	**0.0095**	0.1207	0.3482	0.5515
Credit Spread	-0.1339	0.1145	-1.1688	0.2434	-0.0471	0.1291	0.3172
Term Spread	-0.0149	0.0306	-0.4880	0.6259	-0.0334	0.0144	0.0616
IP Growth	0.0190	0.0076	2.4940	**0.0132**	0.0031	0.0188	0.0231
Real Consuption Growth	0.0020	0.0044	0.4575	0.6476	-0.0054	0.0018	0.0090
Real T-bill Rate	0.0199	0.0300	0.6616	0.5087	-0.0279	0.0187	0.0682
Unexpected Inflation	-0.0206	0.0064	-3.2095	**0.0015**	-0.0211	-0.0148	-0.0007
Bond Risk Factor	-0.0259	0.0719	-0.3605	0.7187	-0.1449	-0.0218	0.0916
Liquidity Factor	0.3172	0.1560	2.0341	**0.0428**	0.0312	0.3214	0.5719
	Bayesian time-varying parameter model (with stochastic volatility)						
Intercept	0.5862	0.0787	7.4482	**0.0000**	0.4575	0.5889	0.7172
Market	0.2197	0.0988	2.2237	**0.0269**	0.0472	0.2220	0.3786
Credit Spread	0.0139	0.0919	0.1516	0.8796	-0.1381	0.0100	0.1710
Term Spread	0.0030	0.0213	0.1401	0.8887	-0.0312	0.0020	0.0382
IP Growth	-0.0079	0.0075	-1.0473	0.2958	-0.0209	-0.0077	0.0036
Real Consuption Growth	0.0047	0.0040	1.1780	0.2397	-0.0013	0.0046	0.0114
Real T-bill Rate	0.0067	0.0216	0.3094	0.7573	-0.0278	0.0055	0.0424
Unexpected Inflation	-0.0092	0.0054	-1.6933	**0.0914**	-0.0183	-0.0090	-0.0001
Bond Risk Factor	-0.0126	0.0492	-0.2550	0.7989	-0.0982	-0.0093	0.0666
Liquidity Factor	0.2071	0.1050	1.9726	**0.0495**	0.0325	0.2066	0.3720
	Bayesian model with time-varying betas (No stochastic volatility)						
Intercept	0.5550	0.2775	1.9996	**0.0464**	0.2500	0.5421	0.8540
Market	0.3006	0.1448	2.0758	**0.0388**	0.0291	0.3028	0.5814
Credit Spread	0.1162	0.1582	0.7346	0.4631	-0.1080	0.0967	0.3963
Term Spread	0.0130	0.0550	0.2368	0.8130	-0.0812	0.0105	0.1120
IP Growth	-0.0067	0.0101	-0.6573	0.5115	-0.0218	-0.0060	0.0111
Real Consuption Growth	0.0030	0.0063	0.4788	0.6324	-0.0072	0.0026	0.0141
Real T-bill Rate	0.0191	0.0498	0.3837	0.7014	-0.0620	0.0165	0.1016
Unexpected Inflation	-0.0024	0.0080	-0.3066	0.7594	-0.0172	-0.0021	0.0103
Bond Risk Factor	0.0431	0.1119	0.3847	0.7007	-0.1242	0.0313	0.2512
Liquidity Factor	0.0474	0.2419	0.1958	0.8449	-0.3394	0.0229	0.4512

Table B.3: Marginal Likelihoods and Bayes Factors Across Alternative Model Specifications

This table reports the values of the marginal log-likelihoods and the relative Bayes Factors for different model specifications. The values reported are also disaggregated by computing the contributions coming from each of the portfolios under investigation. *B-TVB-SV* stands for Bayesian time-varying betas, stochastic volatility model, while *B-TVB* and *B-TVP* are, respectively, the dynamic Bayesian model restricted to have constant conditional volatility and random-walk betas. *Fama-MacBeth* is the standard two-step procedure. *BF1* is the Bayes Factor for the B-TVB-SV model vs. the no-stochastic volatility restriction. Likewise, *BF2* and *BF3* are the Bayes Factors comparing the B-TVB-SV model with the B-TVP and the Fama-MacBeth approaches, respectively.

	B-TVB-SV	**B-TVB**	**B-TVP**	**Fama-MacBeth**	**BF1**	**BF2**	**BF3**
			10 Industry Portfolios, Value-Weighted				
Non Durable Goods	**-445.39**	-1408.71	-635.40	-3131.83	963.32	190.00	2686.44
Durable Goods	**-700.77**	-1980.33	-832.07	-4412.78	1279.56	131.29	3712.01
Manufacturing	**-330.98**	-1199.96	-522.95	-3851.11	868.98	191.97	3520.13
Energy	**-789.61**	-1793.14	-821.64	-2687.83	1003.53	32.03	1898.22
High Tech	**-571.53**	-1732.31	-717.08	-7269.27	1160.78	145.55	6697.74
Telecommunications	**-614.18**	-1634.38	-734.46	-3353.11	1020.20	120.28	2738.93
Shops and Retail	**-481.33**	-1370.61	-648.98	-4271.70	889.29	167.65	3790.38
Health	**-613.64**	-1591.56	-706.14	-3107.84	977.92	92.50	2494.21
Utilities	**-572.88**	-1684.43	-698.85	-1955.89	1111.55	125.96	1383.01
Other	**-270.90**	-1345.87	-519.80	-6041.50	1074.97	248.90	5770.60
			10 Size-sorted Portfolios, Value-Weighted				
Decile 1	**-620.66**	-1756.62	-725.44	-7211.50	1135.96	104.77	6590.84
Decile 2	**-535.44**	-1632.89	-678.94	-6578.42	1097.46	143.50	6042.98
Decile 3	**-428.11**	-1337.08	-616.02	-6506.86	908.96	187.90	6078.74
Decile 4	**-392.01**	-1262.43	-589.11	-7127.84	870.43	197.11	6735.84
Decile 5	**-335.93**	-1134.01	-543.20	-7517.14	798.08	207.27	7181.21
Decile 6	**-259.32**	-932.83	-506.59	-8008.70	673.51	247.27	7749.38
Decile 7	**-202.41**	-923.42	-468.69	-7121.70	721.01	266.28	6919.29
Decile 8	**-149.14**	-882.35	-446.04	-8924.98	733.21	296.90	8775.84
Decile 9	**-95.261**	-601.39	-365.63	-8158.70	506.13	270.37	8063.44
Decile 10	**-54.063**	-544.78	-329.60	-7820.37	490.72	275.55	7766.31
			Bond Returns				
10 - Yrs Treasury	**-188.98**	-1052.55	-422.69	-9201.70	863.57	233.70	9012.72
5 - Yrs Treasury	**-41.972**	-549.00	-320.55	-7951.90	507.03	278.58	7909.93
Baa Corporate Bonds (10-20 years)	**-185.96**	-1050.90	-420.63	-5522.90	864.94	234.67	5336.94
Overall	**-386.11**	-1278.33	-576.98	-5988.50	892.22	190.87	5602.40

Table B.4: Variance Decomposition Tests Across Models

This table reports the results of variance decomposition tests across models. The first two columns show the values from the standard two-step Fama-MacBeth methodology. All rates are in excess of the holding period return on a 1-month T-Bill. VR1 is the ratio of the variance of a model predicted returns and the variance of expected returns estimated from a projection on a set of instruments $\mathbf{Z}_t$. VR2 is the ratio of the variance of the predictable part of returns not explained by a model and the variance of projected returns. The instrumental variables are the lagged monthly dividend yield on the NYSE/AMEX calculated as in Campbell and Beeler (2012), the lagged yield of a Baa corporate bond, and the lagged spread of long- vs. short-term government bond yields. *B-TVB-SV* stands for Bayesian time-varying betas with stochastic volatility, while *B-TVP* and *B-TVB* are, respectively, Bayesian time-varying parameters and time-varying betas models.

	Fama-MacBeth		B-TVB-SV						B-TVP						B-TVB					
	VR1	VR2	VR1			VR2			VR1			VR2			VR1			VR2		
			2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%	2.5%	50%	97.5%
10 Industry Portfolios, Value-Weighted																				
Non Durable Goods	0.368	0.695	0.318	0.454	0.530	0.089	0.144	0.294	0.254	0.320	0.450	0.241	0.328	0.418	0.046	0.217	0.749	0.243	0.719	0.835
Durable Goods	0.499	0.450	0.547	0.833	0.969	0.098	0.131	0.213	0.256	0.376	0.789	0.239	0.678	0.783	0.072	0.464	0.802	0.187	0.519	0.832
Manufacturing	0.717	0.205	0.756	0.916	1.053	0.031	0.173	0.309	0.090	0.453	0.816	0.142	0.418	0.934	0.134	0.492	0.818	0.200	0.569	0.845
Energy	0.698	0.354	0.897	0.987	1.090	0.017	0.056	0.067	0.707	0.832	0.952	0.115	0.221	0.305	0.613	0.865	0.938	0.087	0.148	0.296
High Tech	0.660	0.380	0.869	0.941	1.011	0.108	0.138	0.167	0.345	0.546	0.812	0.256	0.431	0.680	0.337	0.614	0.854	0.105	0.309	0.638
Telecommunications	0.494	0.476	0.730	0.888	0.978	0.109	0.154	0.212	0.117	0.436	0.918	0.073	0.636	0.828	0.176	0.418	0.706	0.326	0.533	0.799
Shops and Retail	0.663	0.336	0.615	0.795	0.913	0.045	0.213	0.378	0.112	0.383	0.901	0.139	0.529	0.922	0.190	0.410	0.832	0.159	0.561	0.790
Health	0.428	0.539	0.612	0.830	0.886	0.003	0.116	0.227	0.202	0.479	0.782	0.247	0.518	0.675	0.134	0.490	0.796	0.144	0.560	0.804
Utilities	0.266	0.705	0.432	0.600	0.690	0.017	0.410	0.760	0.066	0.320	0.869	0.134	0.520	0.910	0.101	0.402	0.708	0.266	0.592	0.871
Other	0.278	0.700	0.431	0.615	0.699	0.152	0.357	0.744	0.231	0.375	0.552	0.428	0.592	0.775	0.240	0.385	0.633	0.311	0.618	0.660
10 Size-Sorted Portfolios, Value-Weighted																				
Decile 1	0.314	0.677	0.225	0.309	0.364	0.281	0.648	0.797	0.242	0.342	0.450	0.552	0.681	0.704	0.015	0.296	0.622	0.305	0.750	0.901
Decile 2	0.731	0.182	0.723	0.882	0.985	0.067	0.159	0.255	0.141	0.467	0.842	0.190	0.564	0.761	0.278	0.456	0.689	0.218	0.473	0.704
Decile 3	0.629	0.323	0.611	0.906	0.954	0.002	0.163	0.316	0.278	0.569	0.976	0.131	0.409	0.889	0.168	0.442	0.706	0.187	0.389	0.783
Decile 4	0.603	0.380	0.542	0.820	0.961	0.043	0.178	0.408	0.270	0.520	0.886	0.234	0.516	0.814	0.239	0.488	0.906	0.079	0.476	0.847
Decile 5	0.538	0.497	0.633	0.877	0.999	0.060	0.112	0.311	0.089	0.465	0.951	0.080	0.473	0.820	0.061	0.261	0.552	0.471	0.750	0.906
Decile 6	0.262	0.679	0.620	0.864	1.012	0.030	0.166	0.348	0.070	0.418	0.824	0.120	0.500	0.899	0.154	0.354	0.737	0.198	0.635	0.891
Decile 7	0.448	0.515	0.723	0.846	0.993	0.013	0.183	0.270	0.102	0.478	0.858	0.144	0.360	0.813	0.165	0.372	0.763	0.247	0.604	0.882
Decile 8	0.367	0.650	0.777	0.855	0.942	0.050	0.111	0.276	0.277	0.552	0.834	0.153	0.361	0.737	0.266	0.501	0.809	0.154	0.491	0.693
Decile 9	0.614	0.367	0.686	0.922	1.025	0.053	0.164	0.369	0.169	0.338	0.795	0.095	0.334	0.732	0.059	0.277	0.717	0.244	0.660	0.949
Decile 10	0.585	0.415	0.622	0.768	0.843	0.039	0.229	0.401	0.313	0.599	0.818	0.205	0.460	0.816	0.103	0.354	0.821	0.232	0.616	0.814
Bond Returns																				
10 - Yrs Treasury	0.707	0.249	0.848	0.962	1.099	0.024	0.068	0.112	0.209	0.449	0.837	0.117	0.456	0.759	0.481	0.674	0.808	0.157	0.364	0.482
5 - Yrs Treasury	0.315	0.763	0.819	0.901	1.074	0.047	0.115	0.219	0.307	0.580	0.856	0.199	0.428	0.731	0.282	0.511	0.723	0.246	0.482	0.880
Baa Corp Bonds (10-20 years)	0.443	0.458	0.260	0.376	0.424	0.452	0.563	0.665	0.351	0.467	0.710	0.279	0.512	0.758	0.056	0.245	0.597	0.399	0.675	0.818

Table B.5: Sources of Risk

This table reports the results of variance decomposition tests for each of the risk factors in the estimated B-TVB-SV model. All rates of return are monthly and in excess of the holding period return on a 1-month T-Bill. VR1 is the ratio of the variance of a model predicted returns and the variance of expected returns estimated from a projection on a set of instruments $\mathbf{Z}_t$. VR2 is the ratio of the variance of the predictable part of returns not explained by a model and the variance of projected returns. The instrumental variables are the lagged monthly dividend yield on the NYSE/AMEX calculated as in Campbell and Beeler (2012), the lagged yield of a Baa corporate bond, and the lagged spread of long- vs. short-term government bond yields. The first two columns report the overall values of VR1 and VR2. From the third column, we report the contributions of each factor in percentage terms. The last column reports the interaction effect.

	VR1	VR2	Market	Credit Risk	Term Spread	IP Growth	Cons. Growth	Real T-Bill	Unexp. Infl.	Bond Factor	Liquidity	Interaction
					10 Industry Portfolios, Value-Weighted							
Non Durable Goods	0.724	0.144	0.6668	0.0885	0.0462	0.1794	0.0128	0.0505	0.0201	0.0906	0.0711	-0.2261
Durable Goods	0.833	0.131	1.6078	0.1387	0.0435	0.0489	0.0119	0.0507	0.2293	0.0816	0.0891	-1.3014
Manufacturing	0.916	0.173	1.1235	0.0308	0.0395	0.0407	0.0433	0.0188	0.0388	0.0782	0.0600	-0.4737
Energy	0.987	0.056	0.4865	0.0400	0.0977	0.5681	0.0638	0.0148	0.1874	0.0278	0.0511	-0.5373
High Tech	0.941	0.138	1.3614	0.0205	0.0294	0.0799	0.0048	0.0210	0.0713	0.0257	0.0555	-0.6694
Telecommunications	0.888	0.154	0.5759	0.0257	0.0312	0.0113	0.0318	0.0142	0.0502	0.0488	0.0158	0.1952
Shops and Retail	0.795	0.213	1.0001	0.0415	0.0437	0.1949	0.0249	0.0438	0.0421	0.0819	0.0548	-0.5276
Health	0.830	0.116	0.4855	0.0364	0.0209	0.0999	0.0542	0.0166	0.0174	0.0430	0.0393	0.1869
Utilities	0.600	0.410	0.2630	0.0672	0.0360	0.2747	0.0304	0.0456	0.6862	0.0905	0.0848	-0.5785
Other	0.615	0.357	0.7390	0.0168	0.0220	0.0152	0.0031	0.0100	0.0272	0.0274	0.0311	0.1081
					10 Size-Sorted Portfolios, Value-Weighted							
Decile 1	0.309	0.648	1.0425	0.0916	0.2413	0.0094	0.3509	0.0554	0.2594	0.0635	0.0412	-1.1551
Decile 2	0.882	0.159	0.9682	0.0414	0.2101	0.0217	0.4064	0.0235	0.1365	0.0377	0.0421	-0.8876
Decile 3	0.906	0.163	1.1883	0.0288	0.1446	0.0422	0.3746	0.0228	0.1030	0.0341	0.0386	-0.9769
Decile 4	0.820	0.178	1.1521	0.0398	0.0738	0.0272	0.3701	0.0265	0.0416	0.0411	0.0396	-0.8119
Decile 5	0.877	0.112	1.2364	0.0332	0.0441	0.0221	0.3737	0.0297	0.0180	0.0333	0.0234	-0.8138
Decile 6	0.864	0.166	1.0965	0.0246	0.0175	0.0078	0.1321	0.0179	0.0077	0.0406	0.0243	-0.3690
Decile 7	0.846	0.183	1.0394	0.0175	0.0217	0.0052	0.0720	0.0137	0.0096	0.0205	0.0138	-0.2134
Decile 8	0.855	0.111	0.9876	0.0140	0.0119	0.0039	0.0308	0.0085	0.0076	0.0291	0.0131	-0.1066
Decile 9	0.922	0.164	0.9574	0.0127	0.0147	0.0083	0.0086	0.0082	0.0187	0.0170	0.0130	-0.0585
Decile 10	0.768	0.229	0.7790	0.0085	0.0111	0.0024	0.0297	0.0069	0.0043	0.0141	0.0088	0.1352
					Bond Returns							
10 - Yrs Treasury	0.962	0.068	0.2148	0.5232	0.1083	0.0202	0.0158	0.0616	0.5011	0.2037	0.0906	-0.7392
5 - Yrs Treasury	0.901	0.115	0.3600	0.5275	0.1165	0.0529	0.0264	0.0921	0.7744	0.3006	0.1541	-1.4044
Baa Corporate Bonds (10-20 years)	0.376	0.563	0.0775	0.3299	0.0489	0.0118	0.0058	0.0210	0.2000	0.0798	0.0373	0.1879

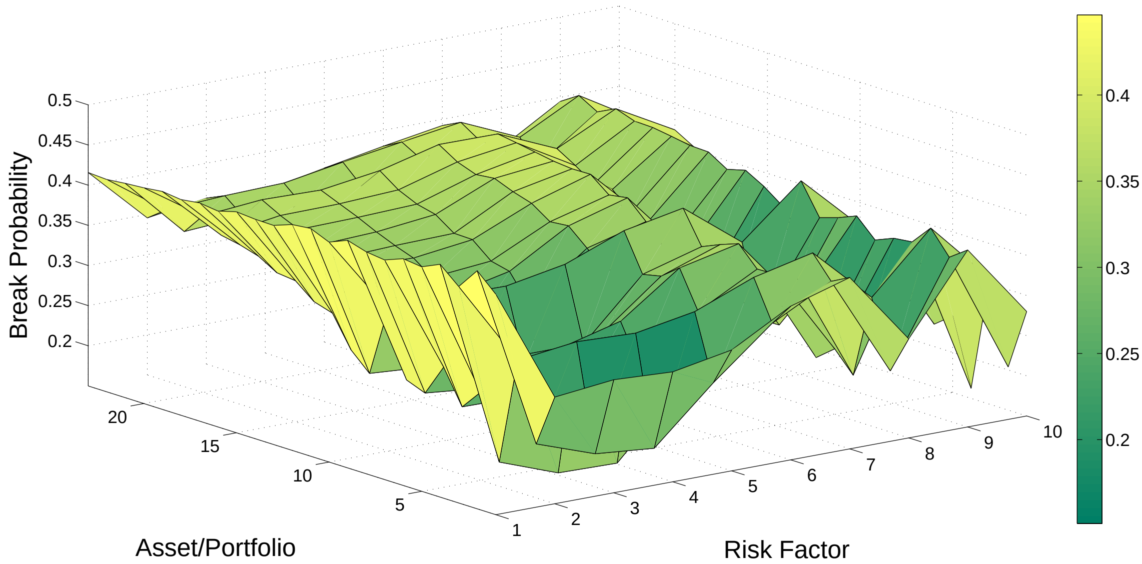
Table B.6: Average Pricing Errors

This table reports the average pricing errors for each of the models under investigation across different subsamples as well as in the full sample. *B-TVB-SV* stands for Bayesian time-varying betas, stochastic volatility model, while *B-TVB* and *B-TVP* are, respectively, the dynamic Bayesian model restricted to have constant conditional volatility and random-walk betas. *Fama-MacBeth* is the standard two-step procedure. The table reports the average (over time), the posterior standard deviation as well as the confidence interval at the 95% level.

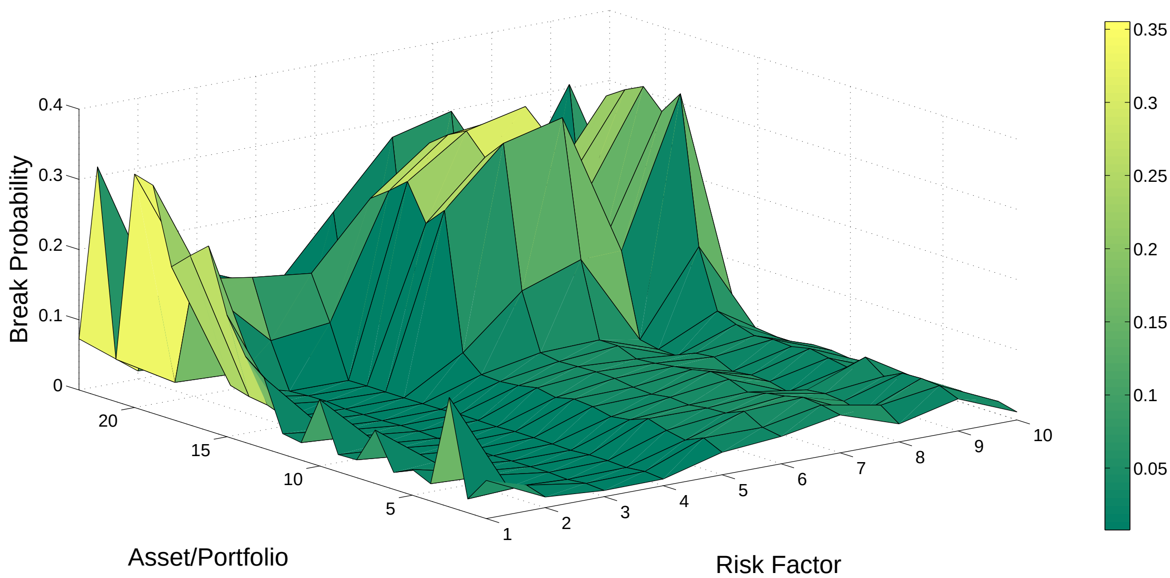
	Average Pricing Errors				
	Mean %	**Std %**	**2.5 %**	**50 %**	**97.5 %**
Panel A: Full-Sample					
B-TVB-SV	**0.2108**	0.0623	0.1363	**0.1902**	0.3231
B-TVB	0.4126	0.1588	0.2512	0.3459	0.7325
B-TVP	0.5401	0.1804	0.3113	0.5061	0.8126
Fama-MacBeth	0.6303	0.0159	0.6107	0.6258	0.6633
Panel B: 1982:01 - 1999:01					
B-TVB-SV	**0.1926**	0.0431	0.1443	**0.1851**	0.2574
B-TVB	0.3935	0.0314	0.3511	0.3921	0.4521
B-TVP	0.5233	0.1653	0.3202	0.4759	0.8101
Fama-MacBeth	0.6278	0.0151	0.6092	0.6238	0.6585
Panel C: 1999:01 - 2011:11					
B-TVB-SV	**0.2624**	0.0672	0.1454	**0.2707**	0.3525
B-TVB	0.4682	0.119	0.2806	0.4501	0.7068
B-TVP	0.6359	0.1993	0.3544	0.6456	0.9027
Fama-MacBeth	0.6354	0.0162	0.6168	0.6321	0.6653
Panel D: 2007:01 - 2011:11					
B-TVB-SV	**0.2891**	0.0613	0.1673	**0.2977**	0.3952
B-TVB	0.5865	0.0906	0.4718	0.5713	0.7559
B-TVP	0.6397	0.1431	0.4029	0.6616	0.8168
Fama-MacBeth	0.6523	0.0179	0.6164	0.6439	0.6799

Figure B.1: Mean Posterior Probability of Breaks in Factor Loadings Across Assets/Portfolios

This figure reports the average over the sample of median posterior probabilities of a break in betas across portfolios and factors for both the B-TVB-SV and B-TVB models. The sample period is 1972:01 - 2011:12. The first 120 monthly observations are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The heating map is reported on the right-hand side.



(a) Bayesian Dynamic Model with Instability in Betas and Conditional Volatility



(b) Bayesian Dynamic model with Constant Conditional Volatility

Figure B.2: B-TVB-SV Factor Loadings: VW Market Portfolio

This figure reports the time series of the posterior median loadings for the market risk factor estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first 120 monthly observations are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The gray area surrounding posterior median plots represents 95% confidence intervals.

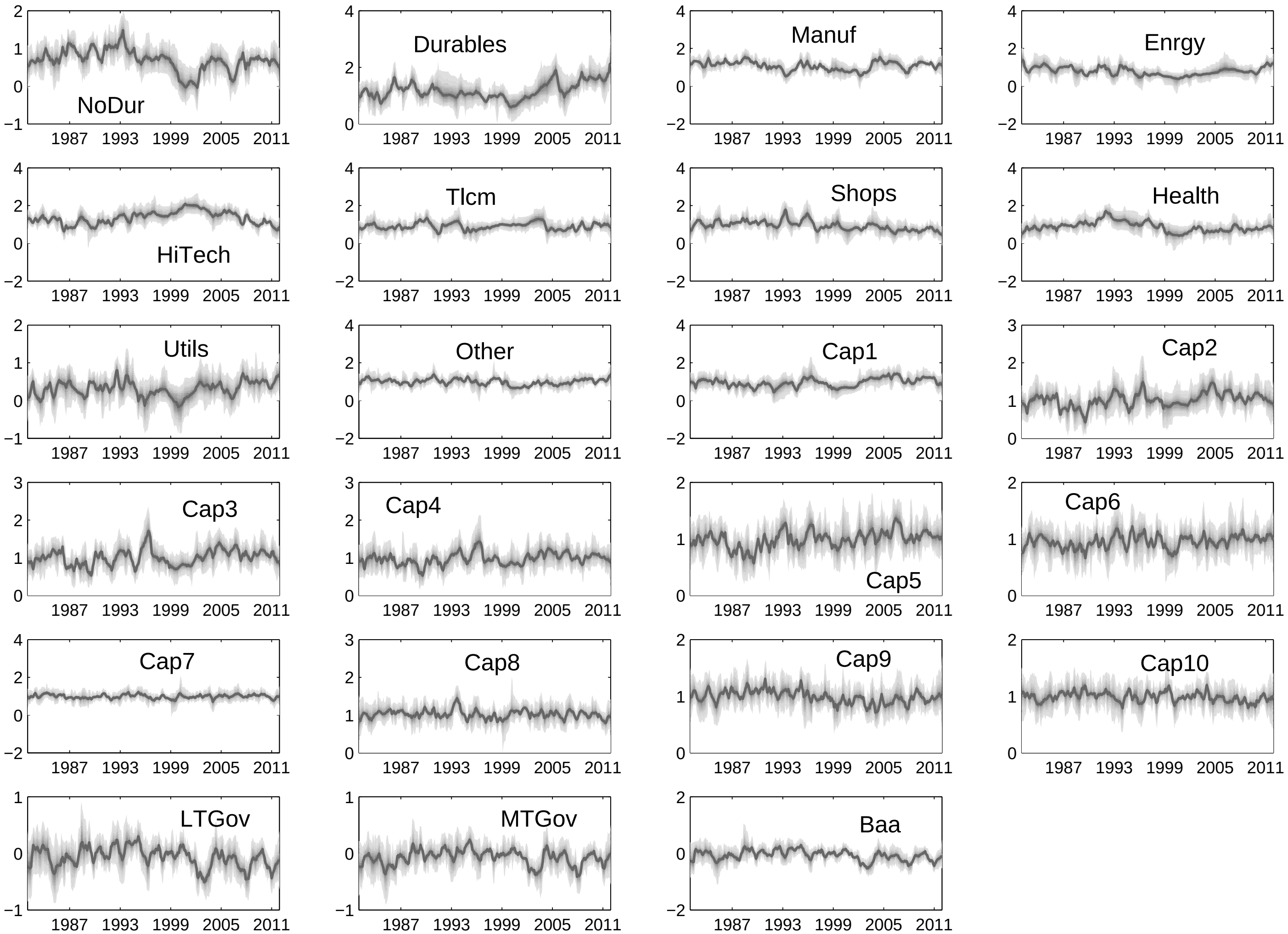


Figure B.3: Factor Loadings Estimated with the Fama-MacBeth Approach: VW Market Portfolio

This figure reports the time series of the of the posterior mean loadings for the market risk factor estimated from a naive 5-year rolling-window estimation approach. The sample period is 1972:01 - 2011:12. The red, dashed lines surrounding posterior mean plots represent 95% frequentist confidence intervals. Asymptotic standard errors are computed assuming absence of cross-sectional dependence among the betas estimates.

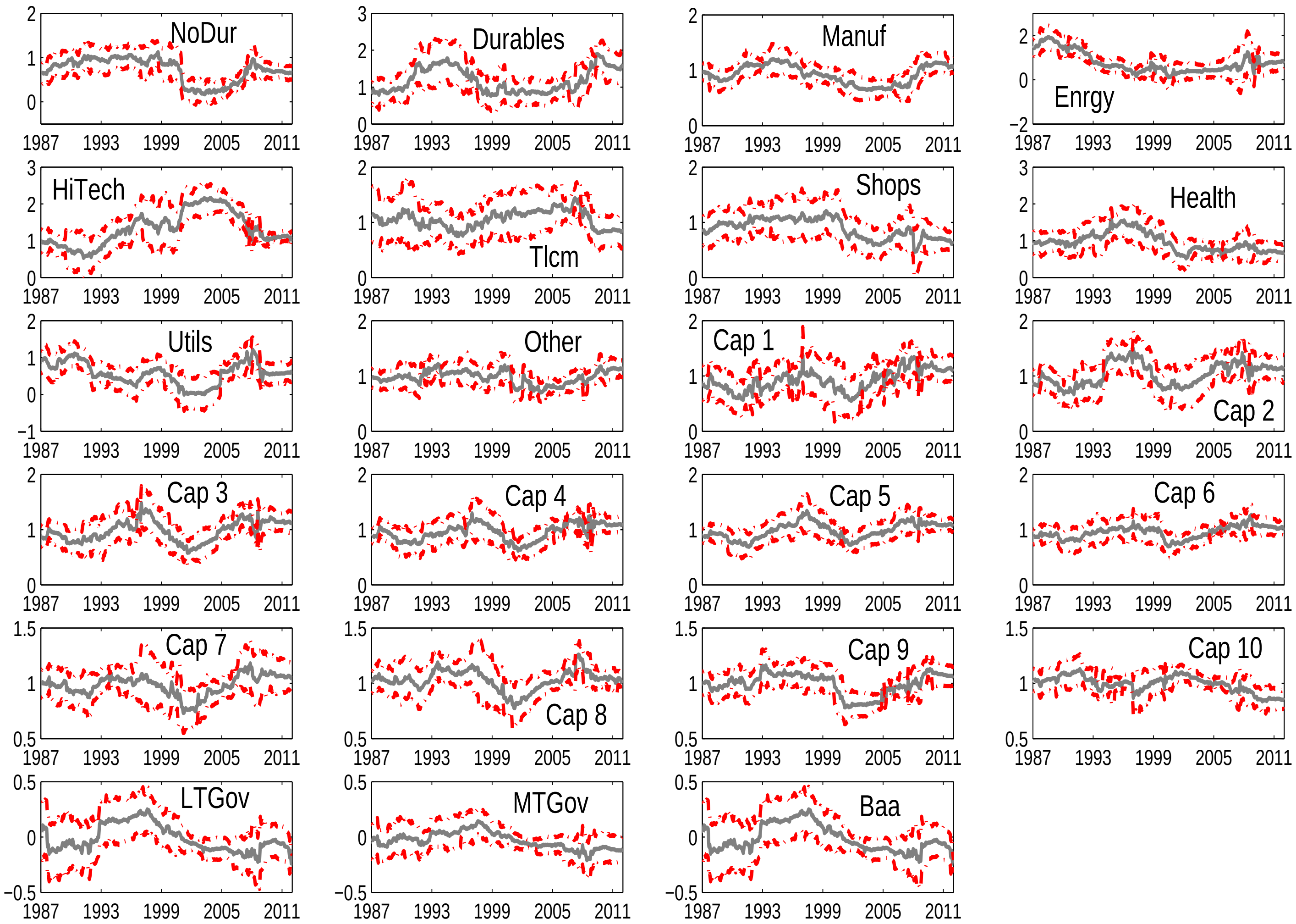


Figure B.4: B-TVB-SV Factor Loadings: Term Spread

This figure reports the time series of the posterior median loadings for the term spread risk factor estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first 120 monthly observations are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The gray area surrounding posterior median plots represents 95% confidence intervals.

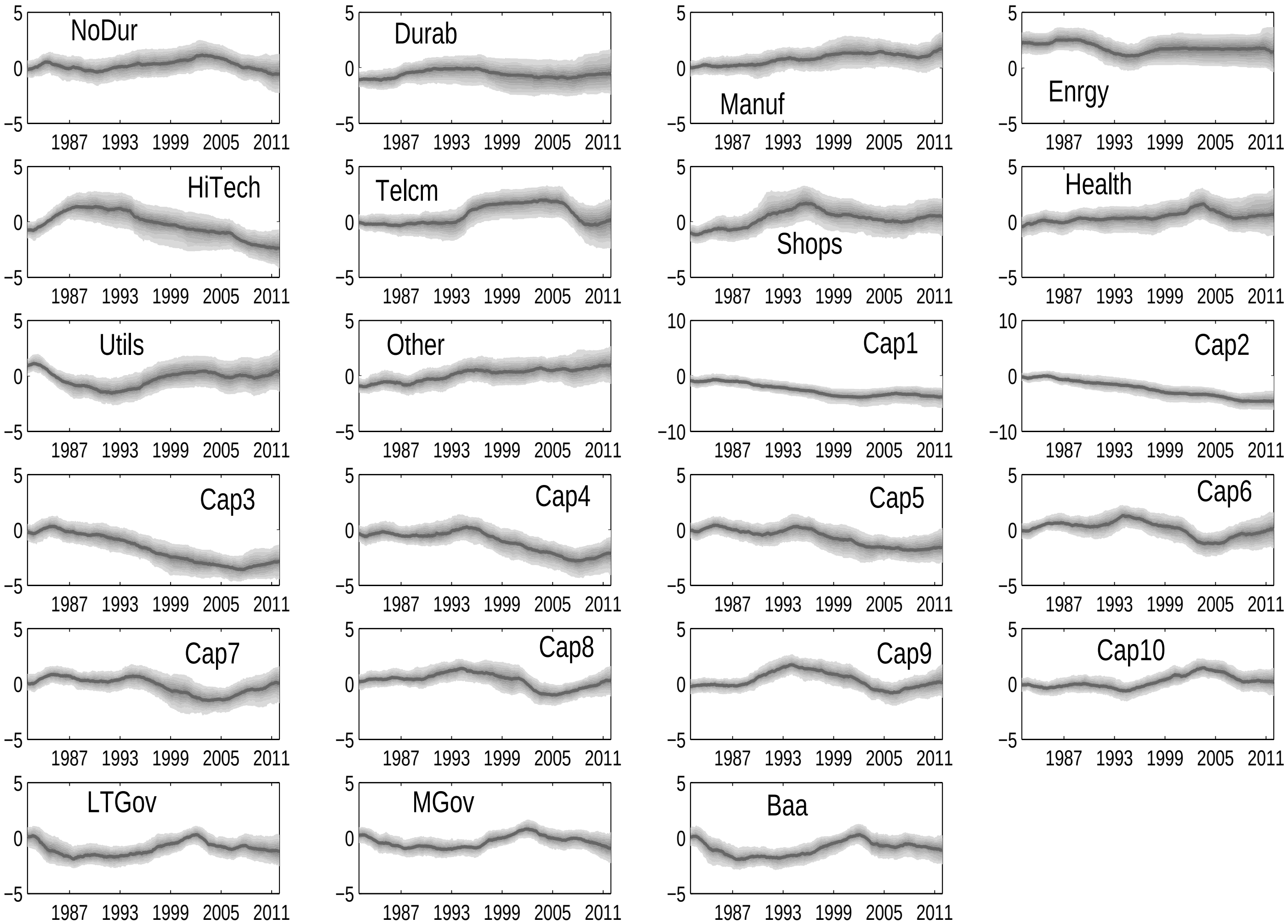


Figure B.5: B-TVB-SV Factor Loadings: Industrial Production

This figure reports the time series of the posterior median loadings for the industrial production growth factor estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first 120 monthly observations are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The gray area surrounding posterior median plots represents 95% confidence intervals.

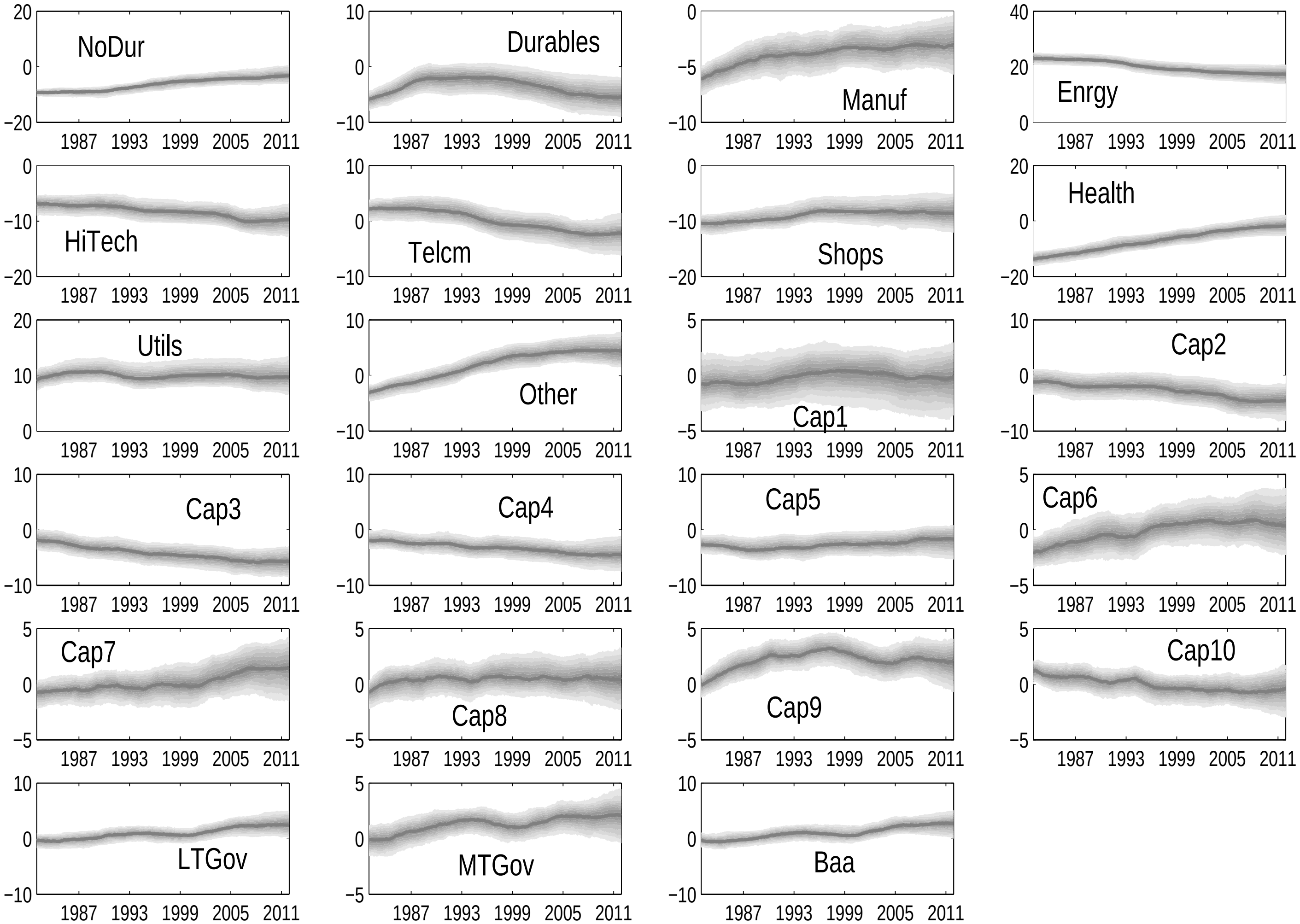


Figure B.6: B-TVB-SV Jensen's Alphas

This figure reports the time series of the posterior medians of the Jensen's alpha from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first 120 monthly observations are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The gray area surrounding posterior median plots represents 95% confidence intervals.

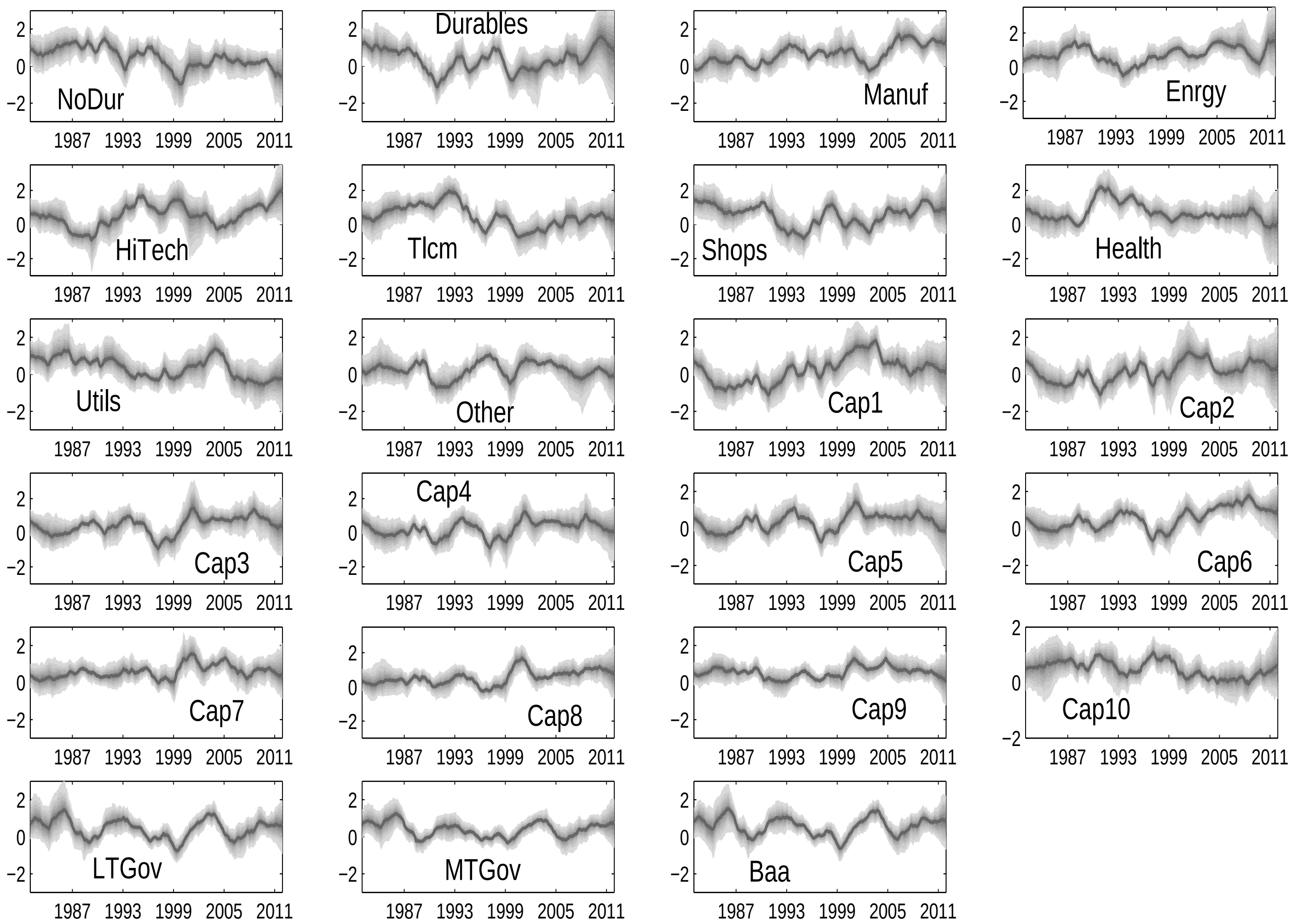


Figure B.7: Jensen's Alphas Estimated with the Fama-MacBeth Approach

This figure reports the time series of the of the posterior means of the Jensen's alphas estimated from a naive 5-year rolling-window estimation approach. The sample period is 1972:01 - 2011:12. The red, dashed lines surrounding posterior mean plots represent 95% confidence intervals.

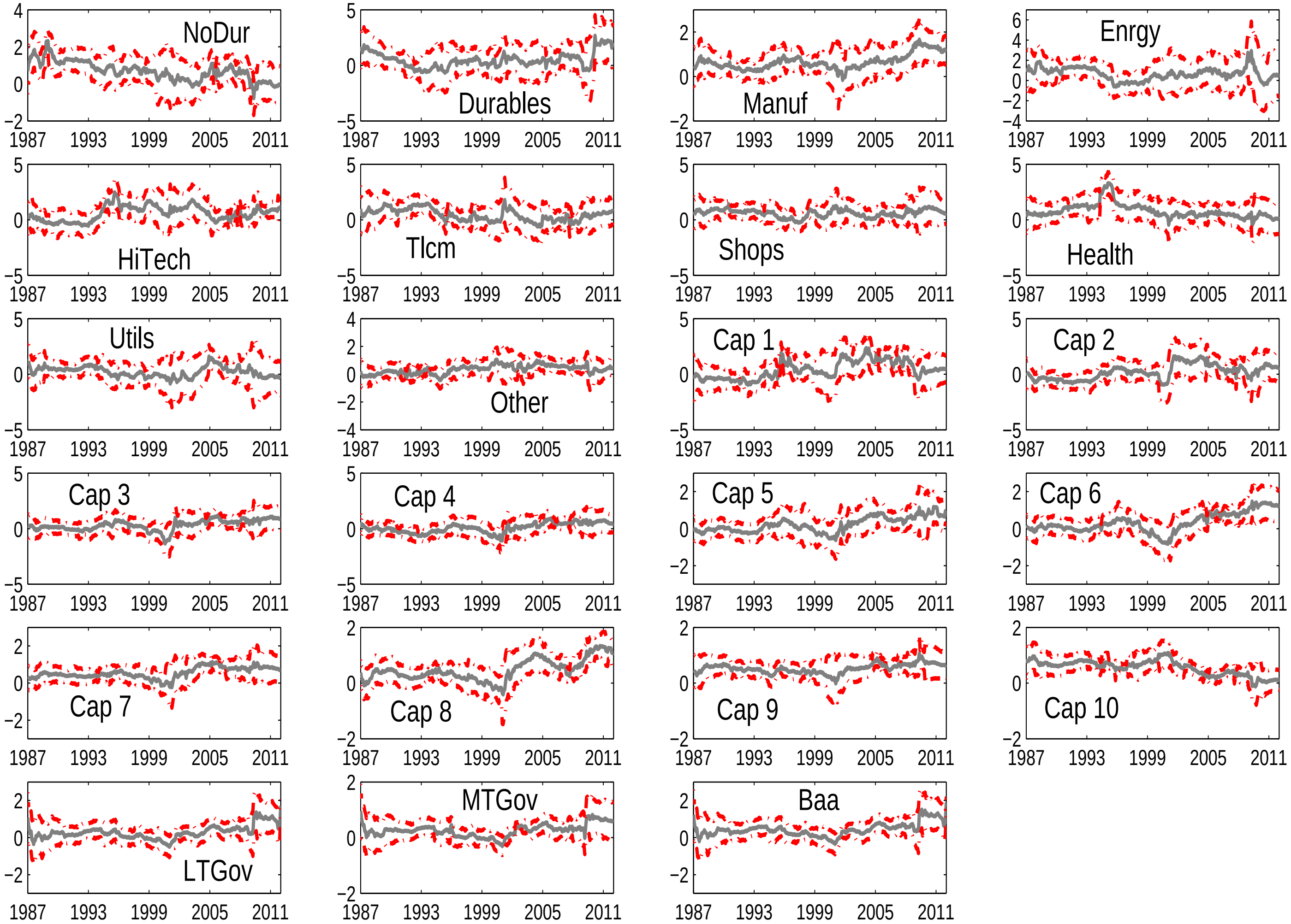


Figure B.8: B-TVB-SV Idiosyncratic Risk Dynamics

This figure reports the time series of the posterior medians for idiosyncratic risk estimated from a dynamic Bayesian model with time-varying betas and idiosyncratic risk. The sample period is 1972:01 - 2011:12. The first 120 monthly observations are used as a training sample in order to calibrate the prior distribution for both latent states and parameters. The gray area surrounding posterior median plots represents 95% confidence intervals.

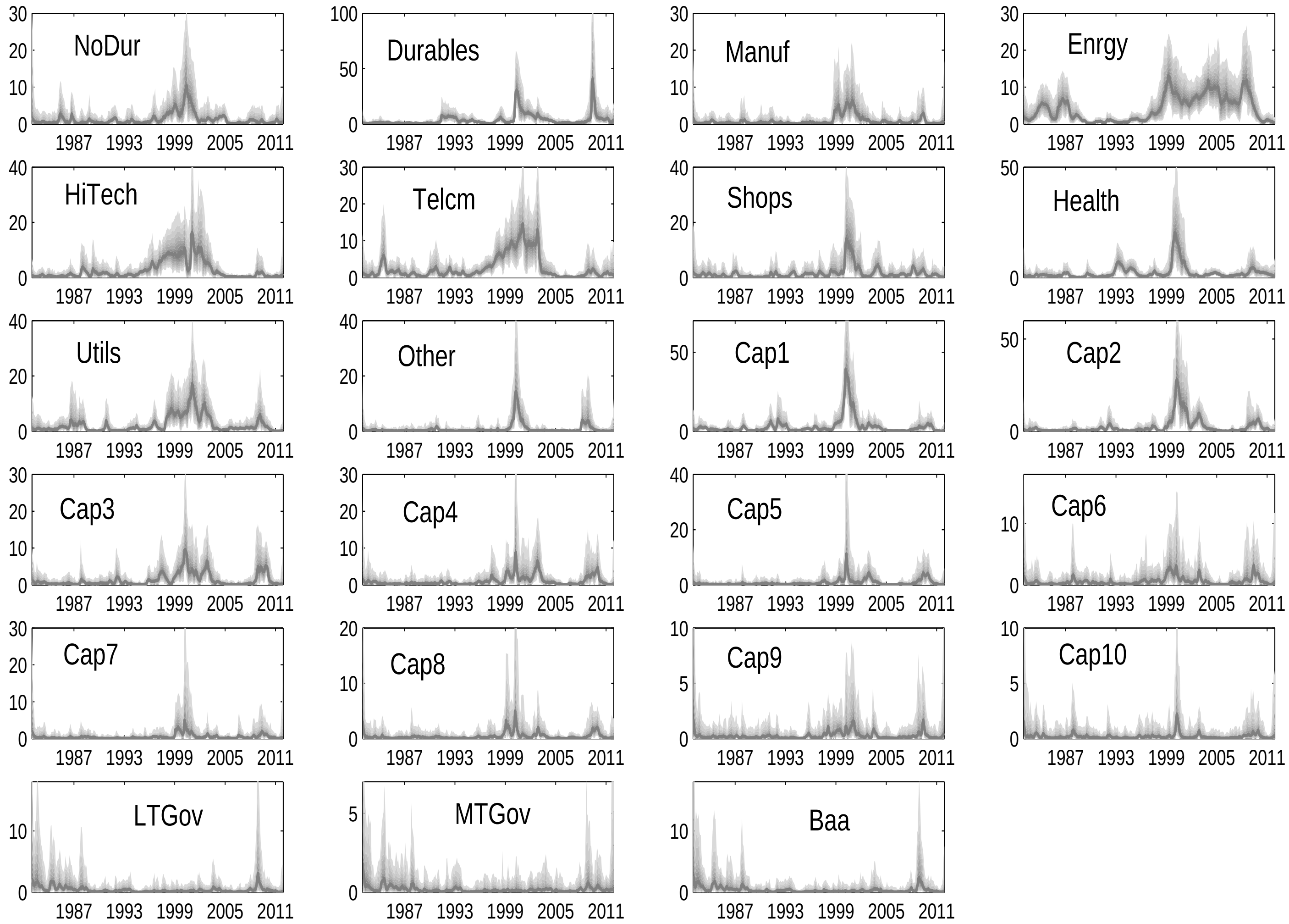
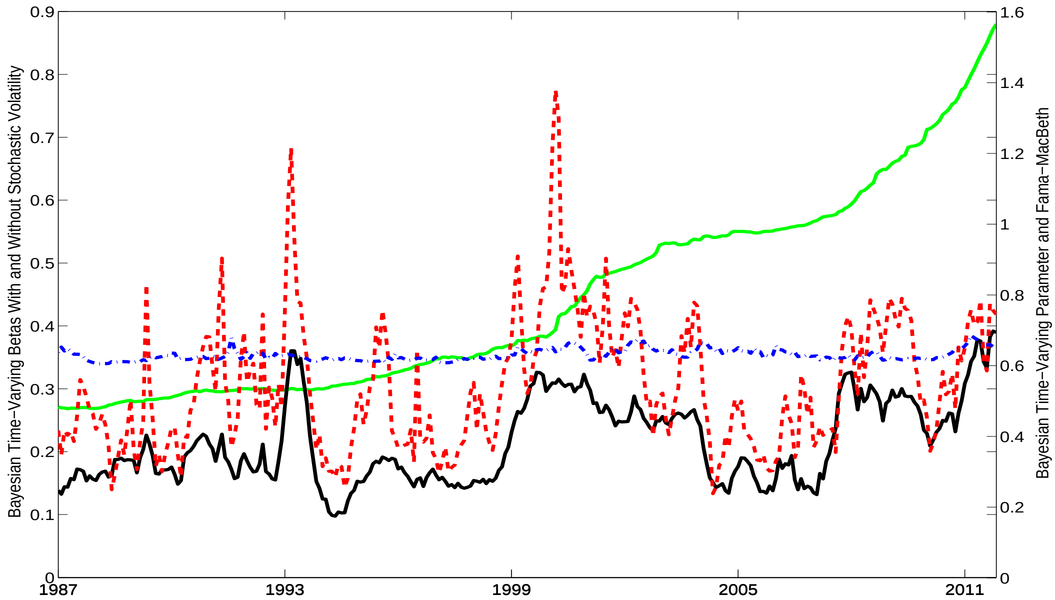
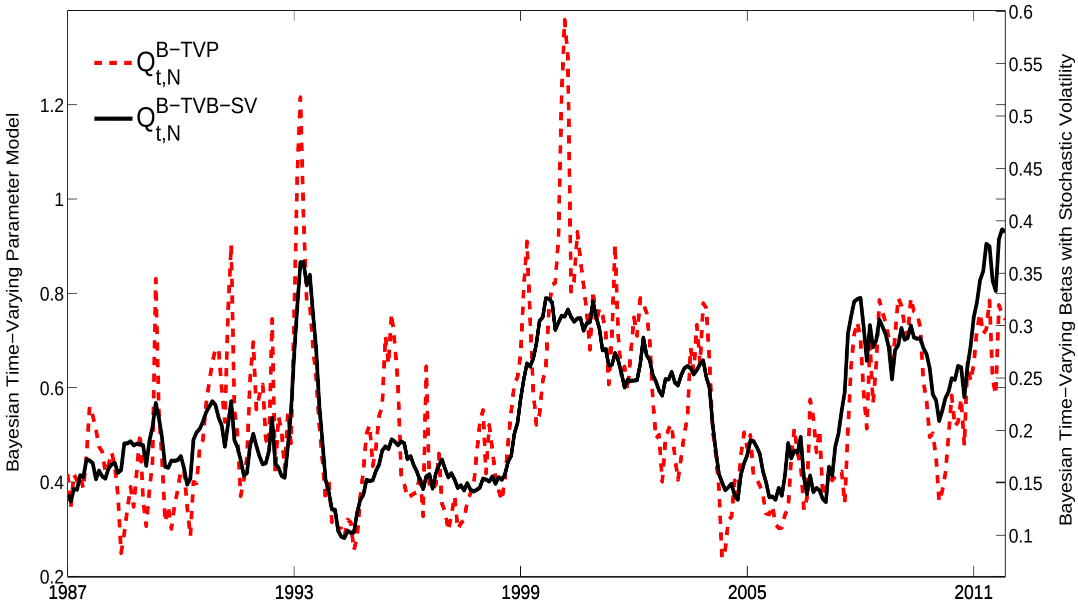


Figure B.9: Average Pricing Errors

This figure reports the time series of the average pricing errors. The sample period is 1972:01 - 2011:12. The first 120 monthly observations are used a training sample in order to calibrate the prior distribution for both latent states and parameters. Panel A reports the average pricing error across models. Panel B reports the rescaled values of the average pricing errors for the B-TVB-SV and the B-TVP models, respectively.



(a) Average Pricing Errors Across Models



(b) Average Pricing Errors for B-TVB-SV and B-TVP (Rescaled)

Chapter 3

Structural Uncertainty and the Higher Moments of Asset Returns

JEL Classification: G12, C11, G17

3.1 Introduction

Conventional wisdom and most of asset pricing research posits an apparent lack of connection between higher moments of market returns and the dynamics of economic fundamentals. Within a general equilibrium setting, smooth and conditionally Gaussian growth rates cannot explain the negative skewness and heavy tails that characterize the unconditional distribution of stock market returns. Such a puzzling disconnect cannot be solved on the basis of mainstream asset pricing models such as the long-run risk setting of Bansal and Yaron (2004) or the habit persistence setting proposed in Campbell and Cochrane (1999), just to cite a few.

The departure from normality in the unconditional distribution of stock market return has been extensively reported in the empirical asset pricing literature "put references and description". Although higher moments in the distribution of asset returns is relatively well established, much less is known about the links with macroeconomic fundamentals. In this chapter, I show that such a disconnect may be explained on the basis of time-varying macroeconomic uncertainty. The baseline assumption is that the average investor does not observe both the state of the economy and the structural parameters governing its dynamics. Incomplete information about the structural parameters play a key role. Indeed, belief updates increase both the impact and the duration of otherwise weak and transitory macroeconomic shocks. In other words, structural uncertainty represents an extra-source of systematic risk which is heavily priced in equilibrium.[1]

The underlying setting is a consumption-based asset pricing model in which a representative agent has incomplete information about the dynamics of economic fundamentals. In the model, the investor has to learn in real-time about both the state of the economy and the structural parameters. The state of the economy has a business cycle frequency and makes the level of economic uncertainty time-varying. In particular, high (low) un-

[1] This is consistent with the argument in **?** (**?**) and Collin-Dufresne et al. (2013). The latter showed that parameter learning may indeed raise a long-run type of risk which is heavily priced under recursive preferences.

certainty typically characterizes periods of economic contraction (expansion), such that macroeconomic uncertainty is negatively related to the business cycle.

Such a rich learning dynamics directly influences the higher moments of the equity premium along two main directions. First, the investor over-weights (under-weights) observations during periods of high (low) economic uncertainty, endogenously generating a time-varying rate of information flow. This asymmetric behavior may vanish asymptotically but is fairly relevant especially in finite samples and with low signal-to-noise ratios, generating, in equilibrium, negative skewness in the unconditional distribution of market returns. Second, the predictive nature of the investor's beliefs increases the perceived riskness of the economic growth, which translates in fat-tails in the unconditional distribution of the market excess returns.

I show that the model is able to replicate higher moments of market returns under conditionally Gaussian economic fundamentals. Both the historical negative skewness and excess kurtosis of the aggregate equity premium fall within the model-implied confidence intervals. The model also may help to solve a set of standard asset pricing puzzles, generating a sizable risk premium under reasonably low risk aversion (Mehra and Prescott 1985), keeping a low level of risk-free rate (Weil 1989) and matching the excess volatility of returns (Shiller 1981). In the model, the log price-dividend ratio is positively in-sample correlated with market excess returns. This generates out-of-sample predictability, which increases with the forecasting horizon (Campbell and Cochrane 1999). The long-term predictive feature is due to the high persistence of the model-implied log price-dividend ratio. This high persistence is due to the permanent nature of belief updates on (the mean posterior estimates of) the structural parameters. I show also that the model generates counter-cyclical conditional volatility of stock returns (? ?) and pro-cyclical valuation ratios (Fama and French 1988a).

I find that states of high macroeconomic uncertainty coincide with well-known economic shocks such as the OPEC I and OPEC II oil crisis, the second gulf war, the dot.com bubble burst and the financial scandals, and the recent great financial crisis. I also provide evidences empirically that parameter learning may represent a risk factor on itself, explaining a significant fraction of the historical excess returns. In fact, the model-implied revision of belief explains an economically relevant fraction of the in-sample excess returns variation. This explanatory power is robust to the inclusion of current and past consumption growth, past equity premium and the belief updates computed from a fixed-parameters model.

3.2 Related Literature

This chapter fits into a growing literature that explores the role of structural learning in self-contained general equilibrium models (Lewellen and Shanken 2002, Weitzman 2007, Johannes et al. 2011 and Collin-Dufresne et al. 2013 among the others). These studies find that, in a consumption-based asset pricing framework, structural learning may help to provide a sensible explanation to some of the standard asset pricing puzzles, which are still at the core of the financial economics literature. **?** (**?**) and Johannes et al. (2011) are the only other chapters exploring real-time parameter learning in a general equilibrium setting with recursive preferences. The former focuses on the time series implications of learning about the dynamics of economic fundamentals and the latter shows, from a broader perspective, the asset pricing implications of parameter learning, especially in terms of the first two unconditional moments.

In contrast my model focuses on higher moments of aggregate equity premium. Lewellen and Shanken (2002) show that parameter uncertainty about the cash-flow dynamics generates predictability and excess volatility, making learning an important tool for testing market efficiency. They considered standard CARA investors and macroeconomic risk is time invariant. Weitzman (2007) show that structural uncertainty may help to match the first two unconditional moments of aggregate returns. However, Bakshi and Skoulakis (2010) points out that the level of structural uncertainty implied by Weitzman (2007) must be implausibly high.

Some of the earlier literature on structural learning in asset pricing goes back to Timmermann (1993), Barsky and De Long (1993) and Timmermann (1996). They focus on the role of learning reduced form pricing kernels. The decision problem (which is the solution of the equilibrium condition) has been substituted out and the role of the agent is limited to forecasting variables which may be typically under her control (e.g. prices) from an asset pricing perspective. In contrast, I focus on the impact of the system of the agent's belief in equilibrium focusing on implications for higher moments of aggregate equity premium.[2]

Finally this chapter is connected to some of other asset pricing literature, such as Veronesi (2000), Veronesi (1999), Brandt et al. (2004), Guidolin and Timmermann (2007) and Lettau et al. (2008), just to cite a few. These chapters show that, by considering state filtering with fixed-parameters may help to explain the standard of asset pricing puzzles and returns predictability. However, these studies mostly assume high persistence of the state variables, because of the inherited amplification mechanism of the exogenous innovations is relatively weak.

3.3 The Asset Pricing Model

3.3.1 The Economy

The underlying setting is a representative agent endowment economy with recursive preferences (Kreps and Porteus 1978, Epstein and Zin 1989, Epstein and Zin 1991 and Weil 1989). These preferences allow to separate the timing of resolution of uncertainty and the relative risk aversion. The wealth-consumption ratio reflects both current fundamentals and the agent's belief about the expected (future) consumption growth rate.

[2]Williams (2003) argues that, fully consider the system of agent's belief may have significant asset pricing implications.

The functional form of life-time utility is

$$V_t = \left\{ (1-\beta) C_t^{1-\frac{1}{\psi}} + \beta \mathcal{R}_t (V_{t+1})^{1-\frac{1}{\psi}} \right\}^{\frac{1}{1-\psi}} \tag{3.1}$$

in which C_t represents consumption at time t, $\psi \neq 0$ the coefficient of intertemporal elasticity of substitution (IES henceforth), $\gamma \neq 1$ the relative risk aversion (RRA henceforth), and β the subjective discount factor. The Stochastic Discount Factor (SDF henceforth) is derived as

$$M_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \left[\frac{V_{t+1}}{\mathcal{R}_t (V_{t+1})} \right]^{\frac{1}{\psi}-\gamma} \tag{3.2}$$

The certainty equivalent $\mathcal{R}_t (V_{t+1})$ collapses to the expected utility operator (Kreps and Porteus 1978).

$$\mathcal{R}_t (V_{t+1}) = \left[E_t \left(V_{t+1}^{1-\gamma} \right) \right]^{\frac{1}{1-\gamma}} \tag{3.3}$$

As pointed out in Epstein and Zin (1989) and Epstein and Zin (1991), the SDF can be rewritten in terms of the wealth-consumption ratio P_t^C

$$M_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\beta \frac{1+P_{t+1}^C}{P_t^C} \right)^{\frac{1/\psi-\gamma}{1-1/\psi}} \tag{3.4}$$

The first component of the pricing kernel is what we would obtain under standard CRRA preferences, while the second component is from the continuation value of the utility function which arises if $\gamma \neq 1/\psi$. Utility maximization is 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 the consumption claim. For each asset i the first order condition takes the form

$$E\left[M_{t,t+1}R_{i,t+1}|y^t\right] = 1 \tag{3.5}$$

with y^t representing the available information at time t, $R_{i,t+1}$ the gross returns on the *ith* asset and $M_{t,t+1}$ the SDF pricing the asset from time t to time $t+1$. Let $\exp\left(\Delta c_{t+1}\right) = \frac{C_{t+1}}{C_t}$, the wealth-consumption ratio is derived by substituting out $R_{c,t+1}$ in (3.5), such that from (3.4) we get

$$P_t^C = E\left[\beta^\rho \exp\left(\left(1-\gamma\right)\Delta c_{t+1}\right)\left(1+P_{t+1}^C\right)^\rho \middle| y^t, \theta\right] \tag{3.6}$$

with $\rho = (1-\gamma)/(1-1/\psi)$, and $P_t^C = P^C\left(z_t\right)$ with z_t the vector of state variables. The wealth-consumption ratio represents the solution of a standard contraction mapping $P^C\left(z_t\right)^{n+1} = T\left(P^C\left(z_t\right)^n\right)$ (see a separate tech appendix for more details). The gross returns on the dividend claim is defined as

$$R_{d,t+1} = \frac{D_{t+1}}{D_t}\frac{1+P_{t+1}^D}{P_t^D}$$

Once $P^C\left(z_t\right)$ is found, the corresponding price-dividend ratio at time t is found similarly as

$$P_t^D = E\left[\beta\exp\left(-\gamma\Delta c_{t+1} + \Delta d_{t+1}\right)\left(\beta\frac{1+P_{t+1}^C}{P_t^C}\right)^{\frac{1/\psi-\gamma}{1-1/\psi}}\left(1+P_{t+1}^D\right)\middle| y^t, \theta\right] \tag{3.7}$$

where $\exp\left(\Delta d_{t+1}\right) = \frac{D_{t+1}}{D_t}$. The real risk-free rate is a direct function of the (conditional) expected SDF $r_{f,t} = -\log\left(E\left[M_{t,t+1}|y^t\right]\right)$. The equity premium is defined as $r_{d,t+1} - r_{f,t+1}$ with $r_{d,t+1} = \log\left(R_{d,t+1}\right)$ and $R_{d,t+1} = \exp\left(\Delta d_{t+1}\right)\left(1+P_{t+1}^D\right)\left(P_t^D\right)^{-1}$. The excess market returns are defined as $r_{d,t+1} - r_{f,t+1}$. Following Bansal and Yaron (2004), Lettau et al. (2008), Johannes et al. (2011), I maintain the assumption that $\gamma > 1$ and $\psi > 1$. As such the agent has preferences for early resolution of uncertainty.[3]

3.3.2 The Dynamics of Fundamentals

The real per-capita consumption growth rate Δc_{t+1} is modeled as a time-varying drift plus noise model (see West and Harrison 1997, Harvey 1981 and Hamilton 1994 for more details)

$$\Delta c_{t+1} = \mu_{t+1} + \sigma_c \epsilon_{c,t+1} \qquad \epsilon_{c,t+1} \sim N(0,1) \tag{3.8}$$

Here σ_c represents the conditional idiosyncratic volatility of the growth rate of consumption. The expected growth rate of consumption, μ_{t+1} evolves as an AR(1) process with ν the corresponding persistence parameter.

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

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

$$\Pi' = \begin{pmatrix} p_{LL} & 1-p_{HH} \\ 1-p_{LL} & p_{HH} \end{pmatrix} \tag{3.9}$$

with

$$p\left(\lambda_{t+1} = H | \lambda_t = H, \theta\right) = p_{HH} \qquad \text{and} \qquad p\left(\lambda_{t+1} = L | \lambda_t = L, \theta\right) = p_{LL} \tag{3.10}$$

[3] The assumption $\psi > 1$ have been kept for comparison purposes. As a matter of fact, even though is still under debate (see Campbell and Beeler 2012) is becoming the standard assumption in the consumption-based asset pricing literature.

I define H (L) as the state of High (Low) Macroeconomic Uncertainty, such that $\sigma^2_{\mu,\lambda_{t+1}} \in \left\{\sigma^2_{\mu,H}, \sigma^2_{\mu,L}\right\}$ with $\sigma^2_{\mu,\lambda_{t+1}=H} > \sigma^2_{\mu,\lambda_{t+1}=L}$.[4] In the transition dynamics p_{HH} (p_{LL}) defines the level of persistence of the high(low)-uncertainty state. The unconditional expected growth rate of both consumption and dividends are not influenced by λ_{t+1}. This makes the model estimates unbiased since the historical expected growth rate of consumption coincides with E_μ.

Following Abel (1999) and Campbell (1986), the aggregate dividend growth Δd_{t+1} is modeled as a rescaled version of the conditional expected growth rate of consumption with $\phi > 1$ the rescaling factor.

$$\Delta d_{t+1} = \mu_d + \phi\left(\mu_{t+1} - E_\mu\right) + \sigma_d \epsilon_{d,t+1} \tag{3.11}$$

The exogenous shocks $[\epsilon_{c,t+1}, \epsilon_{\mu,t+1}, \epsilon_{d,t+1}]' \sim N(0, I_3)$ are independent one among the others. Even though the conditionally volatility of the aggregate dividend is not time varying per se, it inherits time variation when it comes to deal with the predictive distribution. Finally μ_d represents the long-run expected growth rate of the aggregate dividends since $E\left[\mu_{t+1} - E_\mu\right] = 0$.

3.3.3 Learning Dynamics

The standard learning-based asset pricing literature assumes that the individual's may infer an unobservable underlying state given the parameters governing its dynamics are either known or somewhat observable. As such, incomplete information boils down to signal extraction and the posterior belief are simple related to the underlying state of the economy regardless parameter uncertainty (Veronesi 2000,Veronesi 1999, Guidolin and Timmermann 2007, Brandt et al. 2004, Massa and Simonov 2005, Hansen 2007 and Hansen et al. 2008).

[4]This restriction is needed a priori in order to identify the states in the estimation algorithm (see Hamilton 1994 for a related discussion).

In contrast I assume the representative agent does not observe the state of the economy as well as the structural parameters governing its dynamics. The only observable variables are $\Delta c_t, \Delta d_t$. Extending the setting of Weitzman (2007), Bakshi and Skoulakis (2010), and Lewellen and Shanken (2002), as information cumulates the agent simultaneously learns about states z_t and parameters θ, updating her belief via a sequential Bayes' rule. By coupling parameters with states learning I can be more explicit about structural uncertainty.

This learning mechanism is particularly relevant from an asset pricing perspective. In fact, Williams (2003) points out that fully addressing structural uncertainty would make learning to have much more substantial asset pricing implications. Carceles-Poveda and Giannitsarou (2008) shows that signal extraction may not necessarily produce relevant asset pricing implications unless high persistence in the state variables is exogenously imposed. The learning mechanism evolves as in the first chapter on which I refer to for more details, together with the Appendix.

3.4 Model Implications

In the model, the compounding effect of real-time learning and time-varying macroeconomic risk generates negative skewness in the unconditional distribution of the agent's belief on the expected growth rate of consumption. Let $m_t = E\left[\mu_t | y^t\right]$ and $C_t = Var\left[\mu_t | y^t\right]$ be respectively the posterior mean and variance of the expected growth rate of consumption under the agent's filtration (see West and Harrison 1997). The posterior mean and variance once Δc_{t+1} becomes observable evolve as

$$\begin{aligned} E\left[\mu_{t+1} | y^{t+1}\right] &= (1-\nu) E_\mu + \nu m_t + A_{t+1} e_{t+1} \\ Var\left[\mu_{t+1} | y^{t+1}\right] &= \nu^2 C_t + \sigma^2_{\mu,\lambda_{t+1}} + \sigma^2_c \end{aligned}$$

in which $A_{t+1} = R_{t+1} / \left(R_{t+1} + \sigma^2_{\mu,\lambda_{t+1}} \right)$ is the so-called Kalman gain, R_{t+1} the predictive variance of μ_{t+1} and $e_{t+1} = \Delta c_{t+1} - E\left[\Delta c_{t+1} | y^t\right]$ the belief revision at time $t+1$. By construction, $\sigma^2_{\mu,\lambda_{t+1}}$ is positively correlated with the state of macroeconomic risk λ_{t+1}. As such, $\sigma^2_{\mu,\lambda_{t+1}}$ is high (low) when the economy is in a bad (good) state.

Panel A of Figure (C.1) shows the model-implied periods of high macroeconomic risk. These coincide with economic shocks such as Vietnam Buildup 1965, the OPEC I oil crisis 1973 and the Franklin National bank collapse 1974, the OPEC II oil crisis at the end of the 80s, the Gulf War I 1990, and the recession at the early 2000s (dot.com bubble and Enron/Worldcom financial scandals), and the great recession around the 2008/2009.[5]

[Insert Figure (C.1) about here]

In the model, the belief updates are negatively related to the state of economic risk. These updates are not symmetric given the counter-cyclical nature of the signal-to-noise ratio. Panel B shows the relation between e_{t+1} and the state λ_{t+1}. The scatter plot shows a u-shaped asymmetric relation. The higher the probability of being in a state of high macroeconomic risk the more negative is the revision of the agent's belief, with an unconditional negative correlation equal to -0.39.

These asymmetries endogenously generates negative spikes in the belief revision which translates in negative skewness in the unconditional distribution of the expected growth rate consumption. Figure (C.2) shows the agent's belief. The agent's puts more weight on negative shocks in the dynamics of the expected growth rate of consumption.

[Insert Figure (C.2) about here]

In equilibrium, the log price-dividend ratio and the stochastic discount factor are a

[5] These events are for the most part consistent with the uncertainty shocks identified in Bloom 2009.

direct function of the agent's belief. As such, the asymmetric nature of the belief may be reflected in the sample path of the log price-dividend. Figure (C.3) shows the average (across simulations) log price-dividend and the (conditional) expected value of the SDF, $E_t[M_{t,t+1}]$.

[Insert Figure (C.3) about here]

The top panel shows the expected SDF. Because of Bayesian learning, uncertainty about the parameters governing the dynamics of the economy reduces over time. Thus, $E_t[M_{t,t+1}]$ tends to increase over time. Spikes in macroeconomic risk generates jumps-like terms which inflates its unconditional variance.

Bottom panel of Figure (C.3) shows the average log price-dividend ratio across the sample. The asymmetric impact of the regime changes reflects in sudden negative movements of the valuation ratio. The log price-dividend ratio moves then pro-cyclically as pointed out in Fama and French 1988a. Interestingly the big drop of the valuation ratio across the recent great financial crisis is consistent with the data. Intuitively these negative spikes generate negative skewness in the unconditional distribution of market returns. This is confirmed by figure (C.4), which shows the distribution of the agent's belief about the expected growth rate of consumption as opposed to the model-implied (average across simulations) equilibrium returns.

[Insert Figure (C.4) about here]

Drops in the agent's belief about the economic outlook are positively associated with drops in stock market returns. As mentioned before, this relation is not symmetric and generates negative skewness in the unconditional distribution on the returns of the dividend claim.

In the model, as a secondary order effect, time-varying macroeconomic risk generate a state-dependent Kalman gain. This makes learning quantitatively relevant even in large samples.[6] In fact the informativeness of the data about the conditional expected growth rate of the economy fluctuates over time preventing the agent to fully acknowledge the underlying states even in large samples. Figure (C.5) shows the behavior of the sample path of the Kalman gain.

[Insert Figure (C.5) about here]

The Kalman gain is counter-cyclical, meaning is increasing in the conditional volatility $\sigma^2_{\mu,\lambda_{t+1}}$. In fact A_{t+1} is high (low) in a state of high (low) macroeconomic risk. Figure (C.6) shows a comparative static of $\partial A_{t+1}/\partial\sigma^2_{\mu,\lambda_{t+1}} > 0$;

[Insert Figure (C.6) about here]

As such, in the model the Kalman is positively related to the business cycle.

3.5 Data and Empirical Results

I use quarterly data for the period 1948:Q1 - 2012:Q3. This sample represents the longest available for quarterly data on consumption growth which also contains a various episodes of economic crisis. Consumption is composed by the seasonally adjusted quarterly per-capita series on real personal consumption expenditures of non-durables and services. Growth rates are constructed by taking first differences of the corresponding log series. The aggregate stock market returns are from CRSP and the riskless asset (proxied by the 30 days T-Bill return) from Ibbotson. Aggregate dividends are computed as in Campbell and Beeler (2012) and corrected for repurchases as in Bansal et al. (2005). Aggregate

[6] Weitzman (2007) introduced the concept of ergodic learning by showing that the amount of information an agent would need to learn about the structure of the economy is not enough even in large samples. Bakshi and Skoulakis (2010), however, pointed out that the level of structural uncertainty needed in the Weitzman (2007) framework must be implausibly high.

returns, per-capita consumption, aggregate dividends, as well as the nominal risk-free rate are transformed in real terms by using the CPI deflator from the FREDII database of the Federal Reserve Bank of St. Louis.

3.5.1 Cash-Flows Dynamics

Descriptive statistics for both real per capita consumption and aggregate dividends are reported in the first two columns of Table (C.1). Data are aggregated annually. The standard errors are computed by using a non-parametric bootstrap. The annualized real per-capita consumption growth mean is 1.93% (with bootstrap standard error 0.224), and the unconditional volatility is 1.5% (with 0.151 as standard error). The data also show a slightly negative skewness and excess kurtosis. Both consumption and dividends show a low level of autocorrelation.

[Insert Table (C.1) about here]

The model matches the unconditional moments of both consumption and dividend growth rates reasonably well. Both the historical unconditional mean and volatility fall within the model-implied confidence intervals. On average, both the negative skewness and the excess kurtosis are matched. As far as the autocorrelation is concerned the model-implied confidence intervals match the historical values.

3.5.2 Equilibrium Returns

Mehra and Prescott (1985) show that a standard full-information rational expectations model with a representative agent and CRRA preferences, an implausible high level of risk aversion is needed to explain the equity risk premium. Such a disconnect cannot be solved by increasing the level of risk aversion. In fact, Weil (1989) show that high risk aversion generates a counter-factual high level of the real risk-free rate, generating the so-called riskfree-rate puzzle. Shiller (1981) show that actual returns on the market have a variance too high to be justified by relative changes in the fundamentals, unless a implausible high leverage is considered.

I show that real-time learning and time-varying macroeconomic risk with recursive preferences may mitigate these asset pricing puzzles. Specifically, I extend the results in Johannes et al. (2011) and Collin-Dufresne et al. (2013) by focussing on higher moments of stock market returns and learning the dynamics of the aggregate real dividend growth.[7]

The consumption dynamics is not *ex-post* calibrated but *estimated* in real-time. The first 10 years of model-implied returns are removed as burn-in sample in order to get rid of potential impacts of prior specifications. The preference parameters are not calibrated ad-hoc in order to match the unconditional moments of returns. In contrast, I use different settings in terms of RRA and IES and check for the resulting equilibrium returns.

I assume the agent has preferences for early resolution of uncertainty. As such, a positive revision in the agent's belief increases the price-dividend ratio (Bansal and Yaron 2004). The fact that IES is greater than one in the model does not contrast with the critique in Campbell and Beeler (2012). In fact, with parameter learning most of the variation of the real risk-free rate is due to time-varying belief about the expected growth rate of consumption. This, in addition to the low level of predictability of consumption growth under learning, does not generate high EIS in a Hall-type regression framework, albeit $\psi > 1$ in the agent's preferences.

Table (C.2) reports the unconditional moments of the equity premium, the real risk-free rate and the log price-dividend ratio under the data and different model specifications. The historical equity premium is statistically significant and equal to 5.9% on an annual basis. Excess returns variance is around 17% at the same frequency. The historical market returns show negative skewness (-0.74) and leptokurtosis (4.5). The unconditional annual

[7]Learning about the dynamics of dividends introduces a further element of interest. The noise in the dividends dynamics, indeed, potentially makes the filtering exercise misleading. This is overcome by assuming that both consumption and dividends share the same common conditional drift and the agent uniquely uses consumption in order to filter out μ_t.

Sharpe ratio is around 0.3. The log price-dividend ratio is relatively low volatile and highly persistent with a first order autocorrelation coefficient equal to 0.96.

[Insert Table (C.2) about here]

The top panel shows the model-implied asset pricing results under $\gamma = 2, 5$ and $\psi = 2.5$. As we would expect, relative risk aversion plays a relevant role. With a risk aversion equal to two, both aggregate equity premium and market returns variation are relatively underestimated with an expected value equal to 3.98% and 12.62%, respectively. Both kurtosis and (the absolute value of the) skewness are is slightly lower than what I find in the data. The model captures the persistence of log price-dividend ratio. However, the correlation between the data and the model-implied log price-dividend ratio is relatively low (0.4).

Increasing the level of risk aversion to $\gamma = 5$ the model closely matches most of unconditional asset pricing moments. Both the historical aggregate equity premium and the volatility of market returns fall within the 95% confidence intervals. The model-implied equity premium is around 5.6% and its volatility 16.6% on an annual basis. The excess kurtosis is also matched. Even though underestimated the model-implied skewness of excess returns has the same negative sign as in the data. The correlation between the historical log price-dividend ratio and the model implied one increases to 0.65.

The model results are relatively stable across different values of intertemporal elasticity of substitution. Bottom panel of Table (C.2) shows the results with $\gamma = 2, 5$ and $\psi = 1.5$. Both aggregate equity premium and market returns volatility are underestimated with $\gamma = 2$. The skewness (kurtosis) is slightly overestimated (underestimated). The model matches the unconditional Sharpe ratio as well as the persistence of the log price-dividend ratio. However, the correlation between the model-implied and the historical values is low and around 0.4. With $\gamma = 5$ the model performance improves, matching most of the unconditional moments of market returns. A low level of IES, therefore, does not apparently hurts the model performance in a sensible way.

The regime-switching nature of the macroeconomic risk may help to endogenously generate counter-cyclical volatility of market excess returns (? ?). Figure (C.7) shows the conditional variance of aggregate returns computed from the model with parameter learning and $\gamma = 5, \psi = 2.5$.

[Insert Figure (C.7) about here]

The conditional volatility tends to reduce throughout the sample. This is so since, as the uncertainty about the structural parameters resolves, the conditional volatility of the agent's belief diminishes. However, the presence of regime-dependent macroeconomic risk makes the market returns variance relevant also at the very end of the sample. In fact, across the recent great financial crisis the aggregate conditional volatility is increasing despite the amount of information cumulated.

At the outset of the chapter I argue that parameter learning may play a key role in explaining the unconditional behavior of the market excess returns. I test this assumption by (re)computing the same model with fixed parameters. These are calibrated consistently with the end-of-sample estimates of the model with parameter learning. This is done for the sake of comparison with the model with parameter learning. As such, the state variables $z_t = (\mu_t, \lambda_t)$ are assumed to be relatively low persistent. Table (C.3) shows the results. I use the same combination of RRA and IES as in the previous analysis.

[Insert Table (C.3) about here]

Parameter learning seems to play a fairly relevant role as extra source of risk. Both the equity risk premium and the volatility of returns are underestimated across different levels of RRA and IES. The model produces negative skewness in three out of four model specifications. In fact, the asymmetric nature of learning is preserved under the fixed-parameter case. On the other hand no excess returns is replicated whatsoever. This may be due to the fact that with fixed parameters, the (conditional) predictive distribution of economic fundamentals (under which is solved the equilibrium condition) is Gaussian.

The correlation between the model-implied and the log price-dividend ratio drops to 0.25/0.23 with $\gamma = 5$ and 0.08/0.09 with $\gamma = 2$. Interestingly the persistence of the log price-dividend ratio within the model significantly declines. This is due to the low level of persistence imposed on the state variables dynamics by using the end-of-sample estimates under parameter learning as calibration tool (see results in section 6).

3.5.3 Predictability

A number of empirical studies document that the log price-dividend ratio may help to predict the realized excess returns at long horizons (Campbell and Shiller; Campbell and Shiller 1988a; 1988b, Fama and French 1988b, Campbell and Cochrane 1999 and the references therein). I test the null of long-term predictability by running a forecasting regression;

$$r_{m,t+k} - r_{f,t+k} = \alpha_k + \beta_k \ln\left(P_t/D_t\right) + u_{t,t+k} \qquad u_{t+k} \sim N(0,1) \tag{3.12}$$

for different forecasting horizons $k = 1, 4, 8, 12, 24$, with $r_{m,t,t+k} - r_{f,t,t+k}$ the excess market returns from t to $t+k$. The dependent variable are overlapping quarterly returns. The independent variable is the model-implied log price-dividend ratio $\ln\left(P_t/D_t\right)$.[8]

I report the means, medians and 95% percentiles computed from 20.000 simulations

[8] The log price-dividends ratio $\ln\left(P_t/D_t\right)$ is generated as the ratio between the current market price from the model and the sum of the dividends payout for the previous 4 quarters

$$\frac{P_t}{D_t} = \ln\left(\frac{P_t}{\frac{1}{4}\sum_{j=0}^{3} D_{t-j}}\right) \tag{3.13}$$

of the equity premium at each time t. The sample period covers 1948:Q1 - 2012:Q3. The first 10 years of results are cut as burn-in sample. The average t-statistics are corrected for heteroschedasticity and autocorrelation in the residuals (Newey-West).

[Insert Table (C.4) about here]

Panel A in Table (C.4) shows the results computed from the model with real-time structural learning. The price-dividend ratio significantly predicts the equity premium at long horizons. The median adjusted R^2 at $k = 24$ quarters is around 20%. The predictive power (slope) increases (decreases) as the forecasting horizon increases, consistent with the data.

The intuition for why structural learning may generate excess returns predictability is the same as in Timmermann (1993) and Lewellen and Shanken (2002).[9] Updates in the posterior belief about structural parameters affect the demand for risky asset. For instance, a positive (negative) shock in the agent's belief about the conditional expected growth rate of consumption increases (decreases) the price-dividend ratio under preferences for early resolution of uncertainty. This generates positive (in-sample) correlation between the log price-dividend ratio and the realized returns. The mean-reverting nature of the excess returns invert the in-sample correlation out-of-sample. Thus, even though ex-ante no predictability of returns is perceived, ex post the econometrician observes that relatively high past prices (from high fundamentals) are followed by relatively low future returns. Since economic fundamentals are not persistent, *ex-post* predictability is mainly driven by parameter learning. It is important to recognize that the high persistence of the model-implied log price-dividend ratio suggests that caution should be used in interpreting

[9] Lewellen and Shanken (2002) argued: "When investors have imperfect information about expected returns or cash flows, they must learn about the unknown process using whatever information is available, which can be formally modeled using Bayesian analysis. Parameter uncertainty necessarily affects prices at a given point in time, through its impact on investors belief, [...] this learning process can be a source of predictability in a way that differs from other models with rational investors."

these predictability evidences.[10]

In the model there is not a priori persistence in the dynamics of consumption growth. In fact, the level of persistent coming from the model estimates is relatively low, $\nu = 0.46$. This considerably reduces the amount of consumption growth predictability implied by the model. As such, the persistent dynamics of the equilibrium price-dividend ratio does not predict the highly transitory dynamics of consumption growth, albeit a (weakly significant) in-sample correlation.

Panel B of Table (C.4) shows the results by substituting the excess returns in (3.12) with the realized consumption growth. The negative out-of-sample correlation between the log price-dividend ratio is confirmed at the long-run horizon. The average robust t-stat shows that at $k = 24$ there is weak predictability of consumption growth. However, for the most part the slopes coefficients are not statistically significant and the adjusted R^2 does not show relevant predictive power.

Panel C of Table (C.4) shows the results for the aggregate real dividend growth rate. The absence of predictability of the economic fundamentals is confirmed. This is again inherited by the low level of persistence implied by the model with structural learning. The rationale for why aggregate dividends may not be predictable within the model is as for the consumption growth.

The key argument underlying the model-implied excess returns predictability is the high persistence of the log price-dividend ratio. This persistence is endogenously generated by parameter learning. In fact, by calibrating a fixed-parameter model with the end-of-sample parameters estimates, much of the excess returns predictability is lost. Table (C.5) shows the results for the same set of predictive regressions;

[Insert Table (C.5) about here]

[10]Valkanov (2003) showed that the standard asymptotic distribution of the t-statistics in predictive regressions with persistent predictors may lead to mis-specified results. In some cases then standard measures of the goodness of model fit like the adjusted R^2 may be mis-leading.

The average slope parameters are not statistically significant at the 5% confidence level across different forecasting horizons. The average adjusted R^2 is low with a maximum value of around 3% at $k = 24$. The relatively low R^2 is due to the low volatile model-implied log price-dividend ratio. In fact, in the predictive regression framework this low volatility generates a low signal-to-noise ratio. As such, fundamentals does not produces enough time variation to explain highly volatile excess returns. Panel B and C of Table (C.5) show the results for predictive regression on realized consumption and dividend growth.

3.5.4 Robustness Checks

The previous results show that parameter learning may be a relevant determinant of the aggregate equity premium. The idea is that positive changes in the posterior average belief about fixed parameters may generate a strong form of long-run risk which is heavily priced with preferences for early resolution of uncertainty. In fact, positive updates in the agent's belief about the expected growth rate of consumption (macroeconomic risk) increase (decrease) the price-dividend ratio, generating higher (lower) realized returns.

In order to test this assumption I follow Johannes et al. (2011). I regress the historical equity premium on the revision of the agent's belief about both the conditional expected growth of consumption $E\left[\Delta c_{t+1}|y^{t+1}\right] - E\left[\Delta c_{t+1}|y^{t}\right]$ and macroeconomic risk $Var\left[\Delta c_{t+1}|y^{t+1}\right] - Var\left[\Delta c_{t+1}|y^{t}\right]$. These updates only depend on the changing information y^{t+1} and y^t since both parameter and state uncertainty have been integrated out. I do consider also a set of control variables as additional regressors. Past and current real per capita consumption growth are included in order to control for a direct cash-flow effect on the aggregate returns. Past equity premium is included in order to control form mean reversion effects on the returns dynamics. Belief revision from the fixed-parameter model are also included to test for the specific role of parameter learning. Table (C.7) shows the regressions results.

[Insert Table (C.7) about here]

Column 1 shows that a positive shocks in the agent belief on the expected growth rate of consumption is positively related to the equity premium, consistent with the model implications. Column 2 shows that the higher the perceived macroeconomic risk the lower the lower the realized excess returns. Columns 3-5 show that these results hold controlling for both past and current growth rate in consumption, Δc_{t+1} and Δc_t. The statistical strength of the relationship between shocks in the perceived conditional risk and the excess returns reduces as current and lagged growth rates of fundamentals are included.[11] This result means that revision of the agent's belief have an impact on equity premium beyond a direct cash-flow effect, and is equivalent to test for a direct impact of the belief updates on the log price-dividend ratio. Column 6 shows that past excess returns do not reduce the impact of parameter learning on current equity premium. Column 8 and 9 finally test for the role of parameter learning as an extra source of risk. The regression results show that structural uncertainty may play a relevant role beyond the unobservability of the state of the economy.

One may argue that the adjusted R^2 is relatively modest. I follow Cochrane (1999) using **?** (**?**) in order to asses the relationship between the maximum Sharpe ratio attainable under a predictive regression and its R^2.[12] Let SR the unconditional Sharpe ratio from a buy-and-hold strategy, the relationship between the maximum (unconditional) Sharpe ratio attainable under the predictive model SR_{max} and its R^2 is defined as $SR_{max}^2 = SR^2 + \frac{1+SR^2}{1-R^2}R^2$.

[11]This means that by including past and current growth rates in real per capita consumption makes more difficult to distinguish between a direct cash-flow effect and the impact of the belief changes on the macroeconomic risk.

[12]Even though these are not predictive regression, the formation of the predictors depends on lagged information. This allow to map the reported regression in a predictive setting, at least theoretically.

Table (C.7) shows that, even though small, the adjusted R^2 are economically relevant. For instance, by using $E\left[\Delta c_{t+1}|y^{t+1}\right]-E\left[\Delta c_{t+1}|y^{t}\right]$ and $Var\left[\Delta c_{t+1}|y^{t+1}\right]-Var\left[\Delta c_{t+1}|y^{t}\right]$ as regressors the model generates an unconditional maximum Sharpe ratio which is equal to 1.3 times the Sharpe ratio computed from a buy-and-hold strategy.

3.6 Parameter Estimates and Model Assessment

3.6.1 Parameters Estimates

Conditional on the model dynamics (3.8)-(3.11) the agent simultaneously learns about the underlying states $z_t = (\lambda_t, \mu_t)$ and the parameters $\theta = (E_\mu, \nu, \mu_d, \phi, \sigma_c, \sigma_d, p_{LL}, p_{HH}, \sigma_L, \sigma_H)$. In equilibrium, prices are determined by the agent's posterior mean belief of those state and parameters. Posterior means are computed integrating out both parameters and state uncertainty as

$$E\left[\theta|y^t, \mathcal{M}_i\right] = \int \theta p\left(\theta, z_t|y^t, \mathcal{M}_i\right) d\theta dz_t$$

with $\mathcal{M}_i$ the i_{th} model indicator and y^t the available information about both real per capita consumption and aggregate dividend growth. These are marginal parameters belief as uncertainty about the state z_t is integrated out. Likewise the marginal mean belief about z_t is computed as

$$E\left[z_t|y^t, \mathcal{M}_i\right] = \int z_t p\left(\theta, z_t|y^t, \mathcal{M}_i\right) d\theta dz_t$$

Notice that although λ_t is discrete its marginal posterior mean is not integer valued. In fact $E\left[\lambda_t|\mathcal{M}_i, y^t\right]$ is the probability of being in a state of macroeconomic risk under the model. Figure (C.8) shows the marginal posterior mean of the high- (blue line) and low-risk state (red line).

[Insert Figure (C.8) about here]

The model identifies some of the major macroeconomic events across the post-WWII sample period. The agent's belief tend to remain volatile throughout the sample. This is not surprising since the time-varying nature of the data informativeness makes learning ergodic. Figure (C.9) shows the posterior estimates of the conditional expected growth rate of consumption. The black line represents the marginal posterior mean at each time t. The surrounding gray area shows the conditional distribution.

[Insert Figure (C.9) about here]

Large drops in the expected growth rate of consumption are located across the 80s and the recent great financial crisis where the agent expects a 1% decrease in the consumption growth on a quarterly basis. A less pronounced drop is around the first OPEC I crisis (1974/75), with an expected collapse of 2% on an annual basis. Figures (C.10)-(C.11) show the posterior estimates of the persistent parameter, ν and the long-run growth rate, E_μ respectively

[Insert Figure (C.10) and Figure (C.11) about here]

The level of persistence constantly increases from around 0.15 to 0.4 (on average) at the beginning of the 80s. The auto-regressive coefficient remains almost constant till the end of the sample, with a value of 0.45 at the third quarter of 2012. The unconditional growth rate of consumption remains stable across the sample. After an initial increase to 0.65, the posterior mean steadily decreases from the Vietnam buildup till the OPEC II oil crisis. From the early 80s on, the growth prospects remain stable with an end-of-sample estimates equal to 0.5% on a quarterly basis. This coincides with the historical unconditional mean of consumption growth. Finally Figure (C.12) displays estimates of $E\left[\mu_d|\mathcal{M}_i, y^t\right]$.

[Insert Figure (C.12) about here]

The long-run growth prospects of aggregate dividends drop from 0.85% to 0.4% on a quarterly basis. From the OPEC I oil crisis the unconditional expected growth remains mostly stable with a small peak around the 1990 and early after the recent great financial crisis. As for $E\left[E_\mu|\mathcal{M}_i, y^t\right]$ it is quite easy to see a fairly high level of persistence and low volatility in the long-run expected growth rate.

The visual impression given by figures (C.8)-(C.12) is summarized in Table (C.6). The table reports the end-of-sample estimates for each of the model parameters.[13]

[Insert Table (C.6) about here]

The persistence parameter is statistically significant at the end of the sample with a value equal to 0.46. The unconditional expected growth rate of consumption is exactly identified. In fact, the model-implied unconditional mean $E_T E_\mu = 0.499$ coincides with the historical value $E[\Delta c_t] = 0.5001$. The conditional volatility of μ_{t+1} is higher under the state of high macroeconomic risk. This reflects a property of the data since $Var(\Delta c_t|\lambda_t = H) > Var(\Delta c_t|\lambda_t = L)$. The leverage parameter ϕ is relatively low compared to the literature (Bansal et al. 2007 and **? ?**). The low risk state is fairly persistent, $P_{LL} = 0.844$, while the high-risk state is more transitory, $P_{HH} = 0.52$. As such, the ergodic probability of the state of high risk is around $\pi_H = 0.75$ with an expected duration of 6.4 quarters.

[13]The end-of-sample expected values $E_T\left(.\right)$ can be interpreted as the the maximum likelihood estimates.

The low-risk state has a relatively low ergodic probability, $\pi_L = 0.25$ with an expected duration of around 2.2 quarters.

3.6.2 Model Assessment

One may argue that the model dynamics may not necessarily be supported by the data. In particular the presence of two states of macroeconomic risk may be questioned.[14] I test the null of two regimes by sequentially computing the posterior odds for the two-state model as opposed to a single regime. Posterior odds are based on the marginal posterior probability for each model under investigation. Let $\mathcal{M}_i$ the i_{th} model, its marginal posterior probability given the information y^t is computed as

$$p\left(\mathcal{M}_i|y^t\right) = \frac{p\left(y^t|\mathcal{M}_i\right)p(\mathcal{M}_i)}{\sum_{i=1}^{I} p\left(y^t|\mathcal{M}_i\right)p(\mathcal{M}_i)}$$

The marginal likelihood component is calculated recursively by integrating out both states and parameters uncertainty as

$$p\left(y^t|\mathcal{M}_i\right) = p\left(y_t|y^{t-1}, \mathcal{M}_i\right) p\left(y^{t-1}|\mathcal{M}_i\right)$$

where

$$p\left(y_t|y^{t-1}\mathcal{M}_i\right) = \int p\left(y_t|y^{t-1}, \theta, z_{t-1}, \mathcal{M}_i\right) p\left(z_{t-1}, \theta|y^{t-1}\mathcal{M}_i\right) dz_{t-1}d\theta$$

with $z_{t-1} = (\lambda_{t-1}, \mu_{t-1})$ and $\theta = (E_\mu, \nu, \mu_d, \phi, \sigma_c, \sigma_d, p_{LL}, p_{HH}, \sigma_L, \sigma_H)$ (see the appendix). Figure (C.13) shows the sequential posterior odd probability for the single- vs two-regimes

[14] Multiple regimes in the conditional macroeconomic risk are not considered. In fact, the interpretation of a third regime with a unique impact on the conditional volatility and nothing on the conditional mean may be particularly tricky and ad-hoc.

in the dynamics of macroeconomic risk.

[Insert Figure (C.13) about here]

The green area shows the marginal posterior model probability of the model with a two-regimes in the dynamics of macroeconomic risk.[15] The single regime model is quickly rejected by a the agent. The fact that the agent can learn that real per capita consumption growth may have two regimes in its conditional risk is important. This provide a learning based further evidence to standard Markov regime-switching models such as Cecchetti et al. (1999), Brandt et al. (2004), Guidolin and Timmermann (2007), Ju and Miao (2012), Cogley and Sargent (2008), Moore and Schaller (1996), Mehra and Prescott (1985), just to cite a few. I also test the model in an unconditional sense, by comparing the end-of-sample Bayes factors for both the single- and the two-regimes specifications across the models with fixed parameters and structural learning. Table (C.8) shows the end-of-sample marginal likelihoods as well as the Bayes factors.

[Insert Table (C.8) about here]

The data support again the two-state model. The marginal likelihood is equal to -127.190. Interestingly the model with fixed parameters and two regimes is favored by the data against the model with a single regime and parameter learning. This means that the role of time-varying macroeconomic risk may be relevant per se. For the sake of comparison the persistence parameter is calibrated as in Bansal and Yaron (2004).

[15]For simplicity I set the prior model probability $p(\mathcal{M}_i)$ of each model to be 1/2.

3.7 Concluding Remarks

How smooth/conditionally normal economic fundamentals may generate higher moments on market returns? In this chapter I show that real-time learning and time-varying macroeconomic risk may mitigate such a disconnect.

The underlying setting is an otherwise standard single-agent economy with recursive preferences. Real-time learning is modeled amending the standard full-information rational expectations assumption. As such, the agent is uncertain about both the underlying states of the economy and the parameters governing their dynamics. Time-varying macroeconomic risk is modeled as a two-state Markov regime-switching process.

I show that counter-cyclical macroeconomic risk may generate learning asymmetries around the business cycle and a time-varying flow rate of information. In the model learning asymmetries generate negative skewness in the unconditional distribution of stock market returns, while the pro-cyclical behavior of the signal-to-noise ratio makes learning ergodic. I also show that the predictive nature of the agent's belief about the expected growth rate of (future) consumption may help to generate heavy-tails in aggregate market returns. These higher moments effects are generated with conditionally Gaussian economic fundamentals. Coherently with some of the existing literature I finally provide evidences that parameter learning may be seen as an extra source of risk which is heavily priced with preferences for early resolution of uncertainty.

Appendices

Appendix C

C.1 Prior calibration

The prior information is calibrated by using a training sample in order to calibrate location and scale parameters of each of the prior beliefs distributions. Table (C.9) reports the priors calibration results.

[Insert Table (C.9) about here]

I use as a training sample the annual real per capita consumption and aggregate dividend growth rates kindly put available on the Bob Shiller's website[1]. The model is estimated starting with a flat prior. The posterior obtained at the end of the calibration sample is considered as prior for the post world war II testing period. The testing period covers 1948:Q1 - 2012:Q3. In order to ensure a considerable amount of learning through the testing sample, non-informativeness is imposed considering relatively large variances around the prior hyper-parameter estimates. The prior on the transition mechanism has been set to have persistent regimes, i.e. $p_{LL} = p_{HH} = 0.95$.

C.2 Tables and Figures

[1] http://www.econ.yale.edu/ shiller/data.htm

Table C.1: Cash-Flows Descriptives and Model-Implied Dynamics

This table reports the descriptive statistics on the growth rate of real per capita consumption and aggregate dividend. *Data* reports the data -implied statistics and the standard errors from non-parametric bootstrap. *Model* shows the corresponding model-implied mean, median and percentiles from 20.000 simulations at each time t. The sample period covers 1948Q1 - 2012Q3. The first ten years of quarterly results are removed as a burn-in period for the learning procedure.

Moment	**Data**		**Model**				
	Estimate	**(St.Err)**	**Mean**	**(St.Err)**	**5%**	**50%**	**95%**
$\mathbf{E_T}[\mathbf{\Delta c_t}]$	1.927	(0.224)	1.981	(0.182)	1.734	1.964	2.201
$\sigma_\mathbf{T}[\mathbf{\Delta c_t}]$	1.467	(0.151)	1.435	(0.141)	1.317	1.437	1.671
$\mathbf{skew\Delta c_t}$	-0.276	(0.273)	-0.322	(0.308)	-0.732	-0.291	0.239
$\mathbf{kurt\Delta c_t}$	3.432	(0.675)	3.182	(0.593)	2.621	3.386	4.531
$\rho(\mathbf{\Delta c_t}, \mathbf{\Delta c_{t-1}})$	0.198	(0.102)	0.114	(0.045)	0.097	0.118	0.261
$\mathbf{E_T}[\mathbf{\Delta d_t}]$	1.861	(0.830)	1.864	(0.910)	0.374	1.819	3.512
$\sigma_\mathbf{T}[\mathbf{\Delta d_t}]$	7.446	(1.557)	7.501	(0.641)	5.461	7.491	8.571
$\mathbf{skew\Delta d_t}$	0.196	(0.323)	0.310	(0.369)	-0.274	0.305	0.899
$\mathbf{kurt\Delta d_t}$	4.187	(0.906)	4.951	(1.502)	2.928	4.744	6.674
$\rho(\mathbf{\Delta d_t}, \mathbf{\Delta d_{t-1}})$	0.281	(0.104)	0.271	(0.112)	0.104	0.286	0.443

Table C.2: Asset Pricing Moments: Structural Learning

This table reports the unconditional moments of aggregate market returns, the real risk-free rate and the log price-dividend ratio computed from the model with parameter learning. Results are computed on quarterly basis then aggregated annually. $E_T(.)$ denotes the ex-post mean computed at the end of the sample, which covers 1948:Q1 - 2012:Q3. The first ten years of quarterly results are removed as a burn-in period.

Moments	**Data**		$\gamma = 2, \psi = 2.5$					$\gamma = 5, \psi = 2.5$				
	Estimate	**(St.Err)**	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**
Aggregate Equity Premium												
$\mathbf{E_T}(\mathbf{R_{m,t}} - \mathbf{R_{f,t}})$	5.88	(0.02)	3.98	(0.74)	2.75	3.97	5.20	5.60	(0.74)	4.38	5.60	6.81
$\sigma_\mathbf{T}(\mathbf{R_{m,t}} - \mathbf{R_{f,t}})$	17.38	(0.02)	12.62	(0.61)	11.64	12.60	13.74	16.64	(2.48)	12.87	16.17	20.47
skewness	-0.74	(0.32)	-0.47	(0.20)	-0.83	-0.47	-0.15	-0.48	(0.18)	-0.84	-0.50	-0.11
kurtosis	4.51	(1.34)	3.55	(0.53)	2.93	3.56	4.57	4.47	(0.60)	3.75	4.35	5.56
Sharpe Ratio	0.34	(0.14)	0.32	(0.07)	0.22	0.33	0.45	0.47	(0.07)	0.30	0.44	0.57
Log price-dividend ratio												
$\sigma_\mathbf{T}(\mathbf{pd_t})$	0.73	(0.12)	0.26	(0.12)	0.15	0.21	0.45	0.44	(0.12)	0.35	0.41	0.66
$\rho_\mathbf{T}(\mathbf{pd_t}, \mathbf{pd_t^{model}})$			0.44	(0.08)	0.34	0.45	0.52	0.64	(0.06)	0.53	0.65	0.71
$\rho_\mathbf{T}(\mathbf{pd_t}, \mathbf{pd_{t-1}})$	0.96	(0.07)	0.98	(0.02)	0.95	0.97	0.98	0.96	(0.02)	0.95	0.97	0.99
Risk-free rate												
$\mathbf{E_T}(\mathbf{R_{f,t}})$	1.20	(0.42)	2.15	(0.08)	2.03	2.14	2.32	1.81	(0.13)	1.65	1.82	1.94
$\sigma_\mathbf{T}(\mathbf{R_{f,t}})$	1.08	(0.36)	0.51	(0.21)	0.20	0.56	0.87	0.95	(0.34)	0.46	1.02	1.52

Moments	**Data**		$\gamma = 2, \psi = 1.5$					$\gamma = 5, \psi = 1.5$				
	Estimate	**(St.Err)**	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**
Aggregate Equity Premium												
$\mathbf{E_T}(\mathbf{R_{m,t}} - \mathbf{R_{f,t}})$	5.88	(0.02)	3.37	(0.75)	2.14	3.36	4.62	5.40	(0.70)	4.25	5.45	6.75
$\sigma_\mathbf{T}(\mathbf{R_{m,t}} - \mathbf{R_{f,t}})$	17.38	(0.02)	12.71	(0.60)	11.60	12.58	13.62	15.80	(2.57)	11.87	15.19	20.21
skewness	-0.74	(0.32)	-0.32	(0.13)	-0.69	-0.33	-0.03	-0.60	(0.15)	-1.03	-0.59	-0.23
kurtosis	4.51	(1.34)	3.46	(0.59)	2.78	3.35	4.53	4.63	(0.06)	4.55	4.65	4.73
Sharpe Ratio	0.34	(0.14)	0.29	(0.07)	0.18	0.29	0.40	0.38	(0.11)	0.23	0.37	0.59
Log price-dividend ratio												
$\sigma_\mathbf{T}(\mathbf{pd_t})$	0.726		0.12	(0.10)	0.05	0.11	0.28	0.45	(0.11)	0.36	0.43	0.65
$\rho_\mathbf{T}(\mathbf{pd_t}, \mathbf{pd_t^{model}})$			0.39	(0.17)	0.13	0.39	0.55	0.60	(0.04)	0.52	0.61	0.70
$\rho_\mathbf{T}(\mathbf{pd_t}, \mathbf{pd_{t-1}})$	0.962		0.94	(0.05)	0.89	0.95	0.99	0.93	(0.01)	0.93	0.94	0.95
Risk-free rate												
$\mathbf{E_T}(\mathbf{R_{f,t}})$	1.20	(0.42)	2.31	(0.01)	2.29	2.31	2.34	1.32	(0.08)	1.30	1.32	1.33
$\sigma_\mathbf{T}(\mathbf{R_{f,t}})$	1.08	(0.36)	0.33	(0.07)	0.21	0.34	0.45	0.85	(0.10)	0.62	0.88	1.12

Table C.3: Asset Pricing Moments: Fixed Parameters

This table reports the unconditional moments of aggregate market returns, the real risk-free rate and the log price-dividend ratio computed from the fixed-parameter model. The model parameters are calibrated by using the end-of-sample estimates of the model with parameter learning. Results are computed on quarterly basis then aggregated annually. $E_T(.)$ denotes the ex-post mean computed at the end of the sample, which covers 1948:Q1 - 2012:Q3. The first ten years of quarterly results are removed as a burn-in period.

Moments	**Data**		$\boldsymbol{\gamma = 2, \psi = 2.5}$					$\boldsymbol{\gamma = 5, \psi = 2.5}$				
	Estimate	**(St.Err)**	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**
Aggregate Equity Premium												
$\mathbf{E_T(R_{m,t} - R_{f,t})}$	5.88	(0.02)	1.51	(0.68)	0.37	1.51	2.64	1.71	(0.68)	0.60	1.71	2.84
$\boldsymbol{\sigma_T}\mathbf{(R_{m,t} - R_{f,t})}$	17.38	(0.02)	10.01	(0.49)	9.24	10.04	10.87	10.14	(0.49)	9.33	10.13	10.96
skewness	-0.74	(0.32)	-0.20	(0.17)	-0.47	-0.2	0.07	-0.30	(0.13)	-0.58	-0.30	-0.03
kurtosis	4.51	(1.34)	2.97	(0.32)	2.53	2.93	3.56	2.98	(0.33)	2.53	2.95	3.58
Sharpe Ratio	0.34	(0.14)	0.15	(0.07)	0.04	0.15	0.26	0.17	(0.07)	0.06	0.17	0.28
Log price-dividend ratio												
$\boldsymbol{\sigma_T}\mathbf{(pd_t)}$	0.73	(0.12)	0.10	(0.05)	0.07	0.08	0.23	0.09	(0.04)	0.06	0.08	0.21
$\boldsymbol{\rho_T}\mathbf{(pd_t, pd_t^{model})}$			0.08	(0.02)	0.05	0.08	0.12	0.23	(0.01)	0.21	0.24	0.26
$\boldsymbol{\rho_T}\mathbf{(pd_t, pd_{t-1})}$	0.96	(0.07)	0.42	(0.04)	0.36	0.42	0.48	0.41	(0.02)	0.38	0.41	0.45
Risk-free rate												
$\mathbf{E_T(R_{f,t})}$	1.20	(0.42)	1.11	(0.01)	1.10	1.11	1.12	1.51	(0.01)	1.52	1.53	1.54
$\boldsymbol{\sigma_T}\mathbf{(R_{f,t})}$	1.08	(0.36)	0.68	(0.13)	0.45	0.68	0.90	0.15	(0.04)	0.11	0.16	0.21
Moments	**Data**		$\boldsymbol{\gamma = 2, \psi = 1.5}$					$\boldsymbol{\gamma = 5, \psi = 1.5}$				
	Estimate	**(St.Err)**	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**	**Mean**	**(St.Err)**	**2.5%**	**50%**	**97.5%**
Aggregate Equity Premium												
$\mathbf{E_T(R_{m,t} - R_{f,t})}$	5.88	(0.02)	1.47	(0.69)	0.35	1.49	2.64	1.60	(0.68)	0.57	1.61	2.73
$\boldsymbol{\sigma_T}\mathbf{(R_{m,t} - R_{f,t})}$	17.38	(0.02)	10.00	(0.49)	9.20	9.99	10.81	10.09	(0.49)	9.29	10.08	10.91
skewness	-0.74	(0.32)	-0.19	(0.13)	-0.46	-0.19	-0.07	-0.30	(0.09)	-0.48	-0.31	-0.10
kurtosis	4.51	(1.34)	2.97	(0.33)	2.52	2.93	3.56	2.98	(0.33)	2.52	2.94	3.59
Sharpe Ratio	0.34	(0.14)	0.15	(0.07)	0.03	0.15	0.26	0.17	(0.07)	0.06	0.17	0.28
Log price-dividend ratio												
$\boldsymbol{\sigma_T}\mathbf{(pd_t)}$	0.726		0.07	(0.03)	0.04	0.05	0.15	0.07	(0.03)	0.04	0.05	0.14
$\boldsymbol{\rho_T}\mathbf{(pd_t, pd_t^{model})}$			0.09	(0.04)	0.04	0.09	0.13	0.25	(0.01)	0.24	0.26	0.27
$\boldsymbol{\rho_T}\mathbf{(pd_t, pd_{t-1})}$	0.962		0.42	(0.03)	0.36	0.42	0.48	0.44	(0.02)	0.42	0.44	0.46
Risk-free rate												
$\mathbf{E_T(R_{f,t})}$	1.20	(0.42)	1.47	(0.01)	1.46	1.47	1.48	2.32	(0.02)	2.30	2.33	2.36
$\boldsymbol{\sigma_T}\mathbf{(R_{f,t})}$	1.08	(0.36)	0.65	(0.13)	0.44	0.66	0.88	0.85	(0.10)	0.62	0.88	1.12

Figure C.1: Macroeconomic Uncertainty and Belief Revision

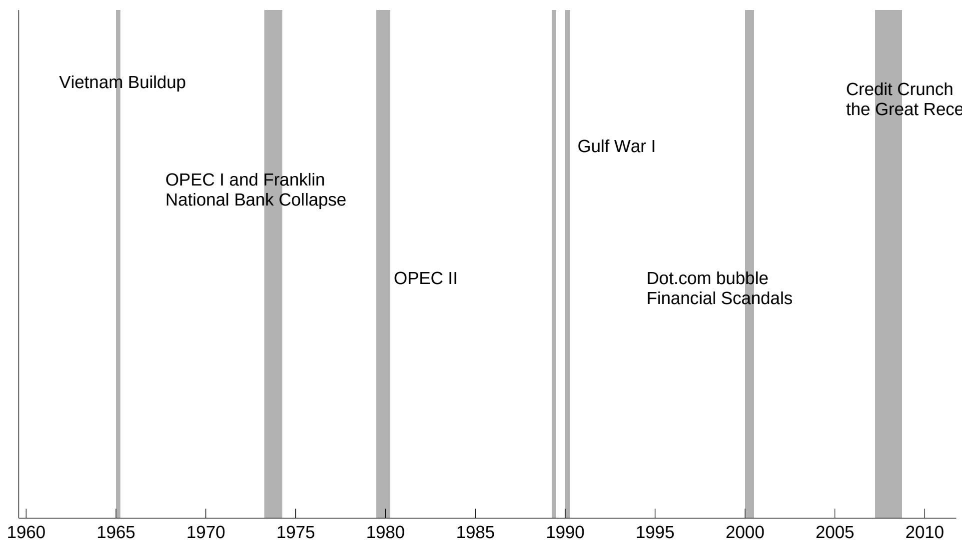


Figure C.2: Density of the Agent's belief About the Expected Growth Rate of Consumption

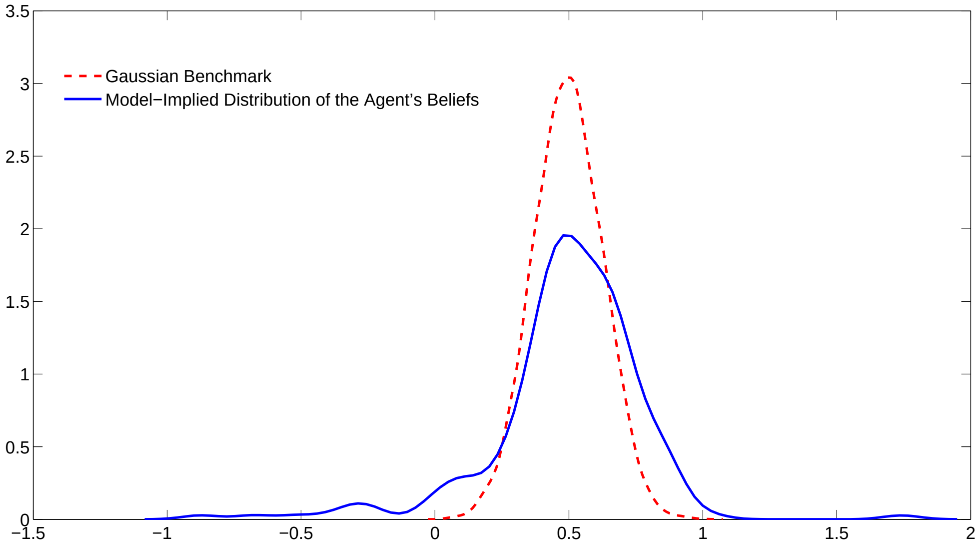


Table C.4: Predictability: Structural Learning

This table reports the results of a set of predictive regressions in which the dependent variable is the aggregate equity premium and the predictor is the model-implied log price-dividend ratio computed from the model with parameter learning. The regressions are computed on overlapping quarterly returns on various forecasting horizons (k quarters from 1 to 24). The t-statistics are adjusted for heteroschedasticity and autocorrelation in the model residuals (Newey-West). * denotes significance at the 10% level. ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The sample covers 1948:Q1 - 2012:Q3. The first ten years of quarterly data are removed as a burn-in period.

Quarters	**Data**			**Model: $\gamma = 5, \psi = 2.5$**				
Panel A: $\sum_{j=1}^{k} (r_{m,t+j} - r_{f,t+j}) = \alpha_k + \beta_k \ln(P_t/D_t) + u_{t,t+k}$								
k	β_k	**(t-stat)**	R^2_{adj}	$\hat{\beta}_k$	$\beta_k(50\%)$	**Avg. t-stat**	$\hat{R}^2_{adj}$	$R^2_{adj}(50\%)$
1	-0.021	(-1.581)	0.010	-0.041	-0.021	(-0.891)	0.004	0.005
4	-0.092	(-1.732)	0.039	-0.152	-0.171	(-2.652)	0.044	0.054
8	-0.144	(-1.823)	0.091	-0.274	-0.243	(-2.674)	0.079	0.096
12	-0.281	(-2.021)	0.192	-0.311	-0.411	(-3.203)	0.124	0.137
24	-0.462	(-4.102)	0.359	-0.690	-0.780	(-3.726)	0.182	0.207
Panel B: $\sum_{j=1}^{k} \Delta c_{t+j} = \alpha_k + \beta_k \ln(P_t/D_t) + u_{t,t+k}$								
k	β_k	**(t-stat)**	R^2_{adj}	$\hat{\beta}_k$	$\beta_k(50\%)$	**Avg. t-stat**	$\hat{R}^2_{adj}$	$R^2_{adj}(50\%)$
1	0.001	(0.311)	0.001	0.002	0.003	(0.572)	0.001	0.001
4	-0.006	(-0.112)	0.004	-0.007	-0.006	(-0.008)	-0.005	-0.004
8	-0.044	(-0.371)	0.005	-0.121	-0.120	(-0.722)	0.005	0.006
12	-0.088	(-0.553)	0.012	-0.251	-0.260	(-0.961)	0.019	0.020
24	-0.023	(-1.081)	0.037	-0.141	-0.142	(-1.651)	0.028	0.029
Panel C: $\sum_{j=1}^{k} \Delta d_{t+j} = \alpha_k + \beta_k \ln(P_t/D_t) + u_{t,t+k}$								
k	β_k	**(t-stat)**	R^2_{adj}	$\hat{\beta}_k$	$\beta_k(50\%)$	**Avg. t-stat**	$\hat{R}^2_{adj}$	$R^2_{adj}(50\%)$
1	0.004	(0.464)	0.004	0.009	0.010	(0.671)	0.002	0.003
4	0.013	(0.426)	0.006	0.069	0.065	(1.011)	0.005	0.005
8	0.020	(0.344)	0.005	0.087	0.083	(1.231)	0.012	0.011
12	0.032	(0.394)	0.008	0.092	0.092	(1.272)	0.008	0.009
24	0.142	(1.454)	0.061	0.101	0.112	(0.791)	0.003	0.003

Table C.5: Predictability: Fixed Parameters

This table reports the results of a set of predictive regressions in which the dependent variable is the aggregate equity premium and the predictor is the model-implied log price-dividend ratio computed from the fixed-parameter model. The model is calibrated by using the end-of-sample estimates of the model with parameter learning. The regressions are computed on overlapping quarterly returns on various forecasting horizons (k quarters from 1 to 24). The t-statistics are adjusted for heteroschedasticity and autocorrelation in the model residuals (Newey-West). * denotes significance at the 10% level. ** denotes significance at the 5% level, and *** denotes significance at the 1% level. The sample covers 1948:Q1 - 2012:Q3. The first ten years of quarterly data are removed as a burn-in period.

Quarters	**Data**			**Model: $\gamma = 5, \psi = 2.5$**				
Panel A: $\sum_{j=1}^{k} (r_{m,t+j} - r_{f,t+j}) = \alpha_k + \beta_k \ln(P_t/D_t) + u_{t,t+k}$								
k	β_k	**(t-stat)**	R^2_{adj}	$\hat{\beta}_k$	$\beta_k(50\%)$	**Avg. t-stat**	$\hat{R}^2_{adj}$	$R^2_{adj}(50\%)$
1	-0.02	(-1.58)	0.010	0.021	0.019	(0.991)	0.007	0.007
4	-0.09	(-1.73)	0.039	-0.055	-0.061	(-0.412)	-0.003	-0.041
8	-0.14	(-1.82)	0.091	-0.127	-0.093	(-0.921)	0.005	0.006
12	-0.28	(-2.02)	0.192	-0.271	-0.274	(-0.843)	0.021	0.020
24	-0.46	(-4.10)	0.359	-0.901	-1.012	(-2.211)	0.034	0.041

Panel B: $\sum_{j=1}^{k} \Delta c_{t+j} = \alpha_k + \beta_k \ln(P_t/D_t) + u_{t,t+k}$								
k	β_k	**(t-stat)**	R^2_{adj}	$\hat{\beta}_k$	$\beta_k(50\%)$	**Avg. t-stat**	$\hat{R}^2_{adj}$	$R^2_{adj}(50\%)$
1	0.001	(0.311)	0.001	0.038	0.039	(0.332)	0.001	0.001
4	-0.006	(-0.112)	0.004	-0.001	-0.002	(-0.023)	-0.001	0.001
8	-0.044	(-0.371)	0.005	-0.021	-0.020	(-0.122)	0.002	0.001
12	-0.088	(-0.553)	0.012	-0.051	-0.060	(-0.611)	0.009	0.010
24	-0.023	(-1.081)	0.037	-0.151	-0.152	(-0.201)	0.005	0.005

Panel C: $\sum_{j=1}^{k} \Delta d_{t+j} = \alpha_k + \beta_k \ln(P_t/D_t) + u_{t,t+k}$								
k	β_k	**(t-stat)**	R^2_{adj}	$\hat{\beta}_k$	$\beta_k(50\%)$	**Avg. t-stat**	$\hat{R}^2_{adj}$	$R^2_{adj}(50\%)$
1	0.004	(0.464)	0.004	0.004	0.010	(0.671)	0.002	0.003
4	0.013	(0.426)	0.006	0.391	0.365	(0.281)	0.012	0.017
8	0.020	(0.344)	0.005	0.508	0.510	(0.821)	0.042	0.041
12	0.032	(0.394)	0.008	1.431	1.230	(0.801)	0.008	0.008
24	0.142	(1.454)	0.061	1.231	1.243	(0.547)	0.004	0.004

Table C.6: Parameters Estimates

This table reports the paramaters estimates of the model

$$\Delta c_{t+1} = \mu_{t+1} + \sigma_c \epsilon_{c,t+1}$$
$$\Delta d_{t+1} = \mu_d + \phi\left(\mu_{t+1} - E_\mu\right) + \sigma_d \epsilon_{d,t+1}$$
$$\mu_{t+1} = (1-\nu)E_\mu + \nu\mu_t + \sigma_{\mu,\lambda_{t+1}}\epsilon_{\mu,t+1} \qquad [\epsilon_{c,t+1}, \epsilon_{\mu,t+1}, \epsilon_{d,t+1}]' \sim N(0, I_3)$$

where $\lambda_t = i$, for $i = H, L$ follows a transition probability

$$p\left(\lambda_{t+1} = H|\lambda_t = H, \theta\right) = p_{HH} \qquad \text{and} \qquad p\left(\lambda_{t+1} = L|\lambda_t = L, \theta\right) = p_{LL}$$

$E_T(.)$ denotes the ex-post mean computed at the end of the testing sample. The period covers 1948:Q1 - 2012:Q3. The first ten years of quarterly estimates are removed as a burn-in sample. Standard Errors are reported in parenthesis.

Panel A: End-of-Sample Estimates					
Parameter	**Model Estimates**				
	Mean	**(St.Err)**	**2.5%**	**50%**	**97.5%**
ν	0.462	(0.057)	0.373	0.464	0.544
$\mathbf{E}_\mu$	0.499	(0.045)	0.434	0.499	0.564
$\sigma^{\mathbf{2}}_{\mu,\lambda_\mathbf{t}=\mathbf{L}}$	0.073	(0.011)	0.053	0.073	0.091
$\sigma^{\mathbf{2}}_{\mu,\lambda_\mathbf{t}=\mathbf{H}}$	0.379	(0.142)	0.179	0.365	0.635
$\mu_\mathbf{d}$	0.483	(0.150)	0.235	0.484	0.727
ϕ	1.406	(0.651)	0.221	1.404	2.624
$\mathbf{p_{HH}}$	0.512	(0.005)	0.511	0.513	0.515
$\mathbf{p_{LL}}$	0.844	(0.002)	0.843	0.844	0.845
$\pi_\mathbf{L}$	0.746	(0.003)	0.745	0.746	0.747
$\pi_\mathbf{H}$	0.254	(0.004)	0.253	0.254	0.256
$\mathbf{Dur_L}$	6.421	(0.001)	6.405	6.422	6.438
$\mathbf{Dur_H}$	2.189	(0.002)	2.184	2.189	2.194
Panel B: Conditional Beliefs					
	Mean	**(St.Err)**	**2.5%**	**50%**	**97.5%**
$\mathbf{E_t}[\mu_\mathbf{t+1}]$	0.500	(0.128)	0.293	0.500	0.705
$\mathbf{Std_t}[\mu_\mathbf{t+1}]$	0.292	(0.050)	0.324	0.285	0.309

Table C.7: Model Testing: Belief Updates and the Equity Premium

This table shows the results from the regression of the belief updates about both the conditional expected growth rate of consumption and macroeconomic risk, on the realized aggregate stock market excess returns. The agent's expectations are computed integrating out both parameter and states uncertainty. The set of control variables are contemporaneous and lagged consumption, past values of the market returns and the belief updates computed from the fixed-parameter model. *SR Max/ SR* is the maximum sharpe ratio attainable under the k_{th} model as a percentage of the unconditional Sharpe ratio of the aggregate market portfolio. The sample period covers 1948:Q1 - 2012:Q3. Robust t-stats corrected for Heteroschedasticity and Autocorrelation (Newey-West) 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 ten years are removed as a burn-in sample.

	Models								
Indipendent: $\mathbf{r_{m,t+1} - r_{f,t+1}}$	1	2	3	4	5	6	7	8	9
$\mathbf{E[\Delta c_{t+1} \| y^{t+1}] - E[\Delta c_{t+1} \| y^{t}]}$	0.246***	0.221***	0.121**	0.216***	0.272**	0.218***	0.283*		0.194***
	(3.001)	(3.132)	(1.981)	(2.771)	(1.991)	(3.198)	(1.867)		(2.914)
$\mathbf{Var[\Delta c_{t+1} \| y^{t+1}] - Var[\Delta c_{t+1} \| y^{t}]}$		−0.091**	−0.089*	−0.031*	−0.084*	−0.092**	−0.074*		−0.088**
		(-1.971)	(-1.892)	(-1.861)	(-1.891)	(-1.962)	(-1.791)		(-1.826)
Controls:									
$\mathbf{\Delta c_{t+1}}$			0.189***		0.231*		0.221***		-0.108
			(2.759)		(1.896)		(3.103)		(-1.359)
$\mathbf{\Delta c_{t}}$				0.124*	-0.631		-0.581		-0.021
				(1.887)	(-0.751)		(-1.482)		(-0.348)
Fixed Parameters								0.202***	0.161***
								(2.818)	(2.410)
$\mathbf{r_{m,t} - r_{f,t}}$						0.021	-0.007		
						(0.408)	(-0.125)		
$\mathbf{R}^2_{adj}$	6.06	6.88	8.55	7.71	8.96	6.93	9.99	5.11	9.41
SR Max/SR	125%	129%	136%	132%	137%	129%	141%	122%	140%

Table C.8: Marginal Likelihoods and Bayes Factors

This table reports end-of-sample marginal likelihoods and corresponding Bayes factors testing the single- vs two-regimes in the dynamics of macroeconomic risk. The table show the results for both the models with parameter learning and fixed parameters. The Bayes factor is computed by considering the two-regimes model under structural uncertainty as the benchmarking model. Here, > 100 means that the Bayes factor of the benchmark against the benchmark is greater than 100. The sample period covers 1948Q1 - 2012Q3. The first ten years of quarterly results are removed as a burn-in period for the learning procedure.

	Structural Learning		**Fixed Parameters**	
	One Regime	**Two Regimes**	**One Regime**	**Two Regimes**
Marginal Likelihood	-138.101	**-127.190**	-143.832	-137.213
Bayes Factor	> 100		> 100	> 100

Table C.9: Prior Calibration

This table reports the hyper-parameters used to calibrate the prior belief. *Location* represents the location parameter. *Scale* identifies the scale parameter. The priors are calibrated on a pre-sample period by using the annual real consumption and aggregate dividend growth from the Robert Shiller's dataset. The scale parameter is set to let the prior specification be uninformative.

Parameter	**Location**	**Scale**
ν	0.975	0.50
$\mathbf{E}_{\mu}$	0.486	0.50
$\sigma^{\mathbf{2}}_{\mathbf{c}}$	0.163	0.50
$\sigma^{\mathbf{2}}_{\lambda_{\mathbf{t}}=\mathbf{L}}$	0.534	0.50
$\sigma^{\mathbf{2}}_{\lambda_{\mathbf{t}}=\mathbf{H}}$	0.137	0.05

Figure C.3: Expected SDF and the Log Price-Dividend Ratio

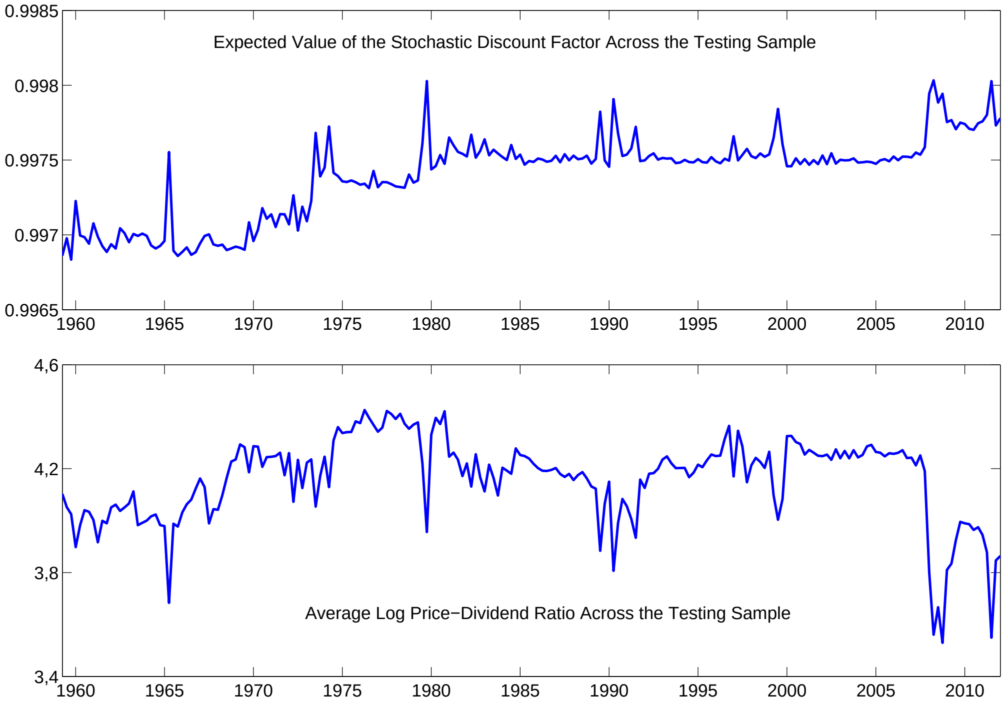


Figure C.4: Agent's Belief and the Model-Implied Excess Returns

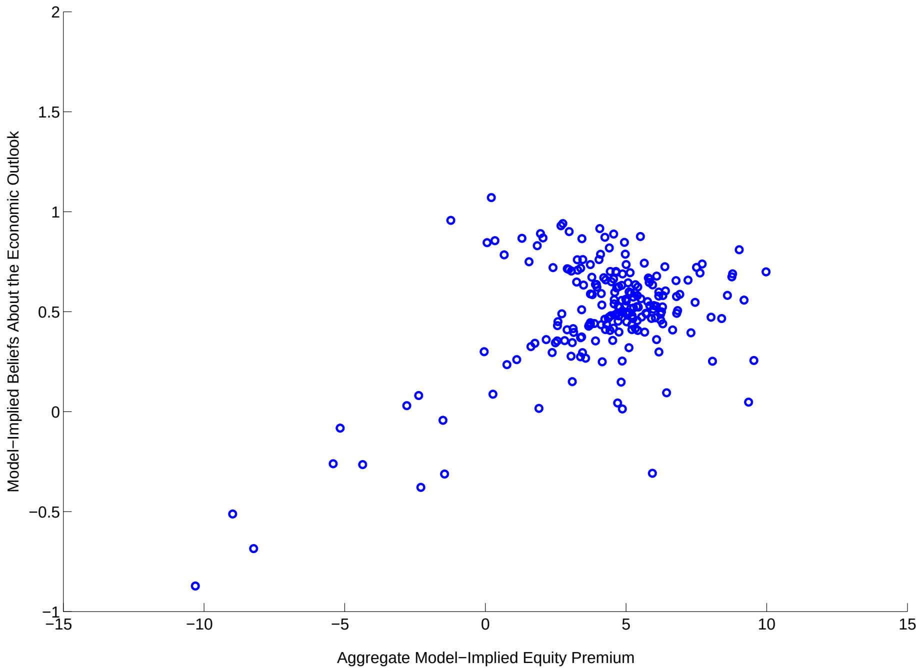


Figure C.5: Data Informativeness

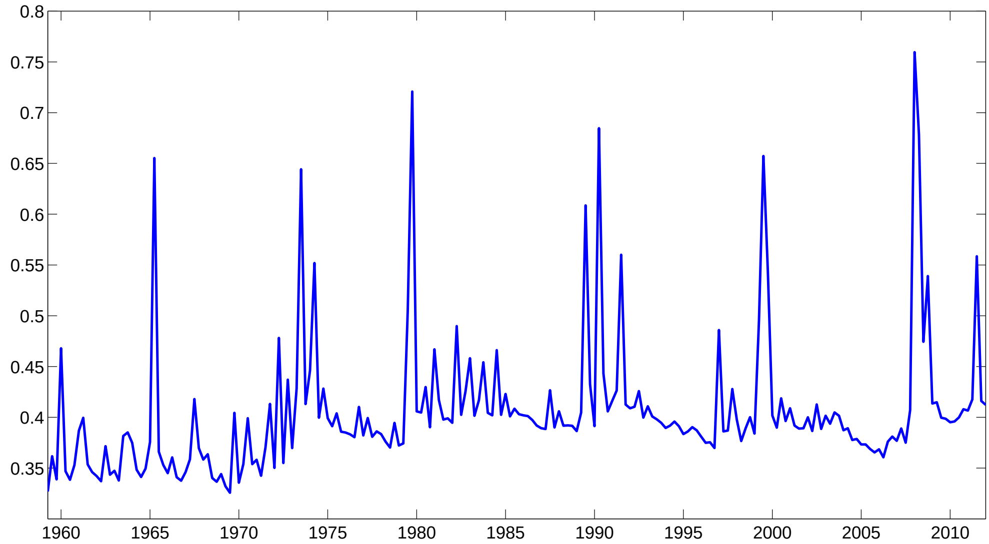


Figure C.6: State Dependence of the Kalman Gain

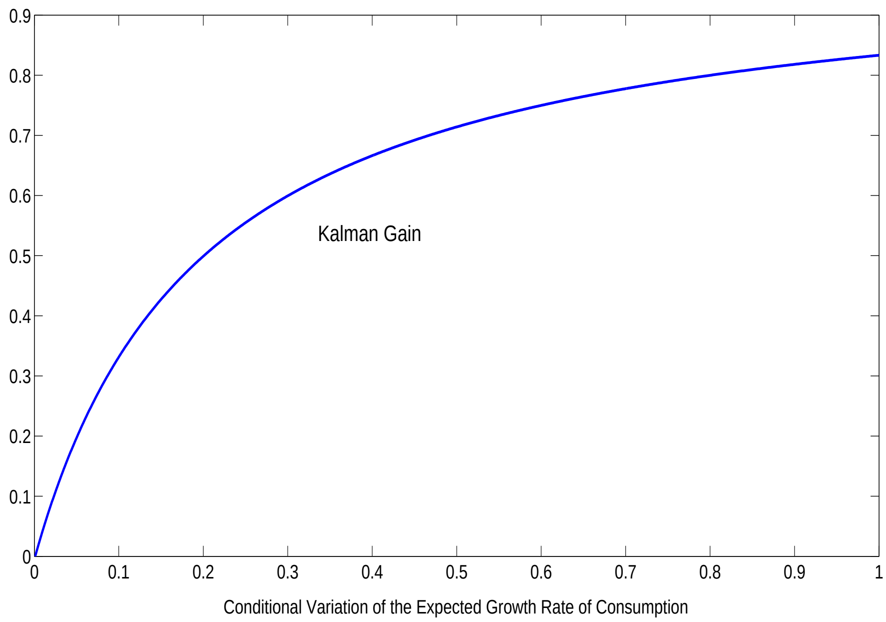


Figure C.7: Conditional Variation of the Returns across the Testing Sample

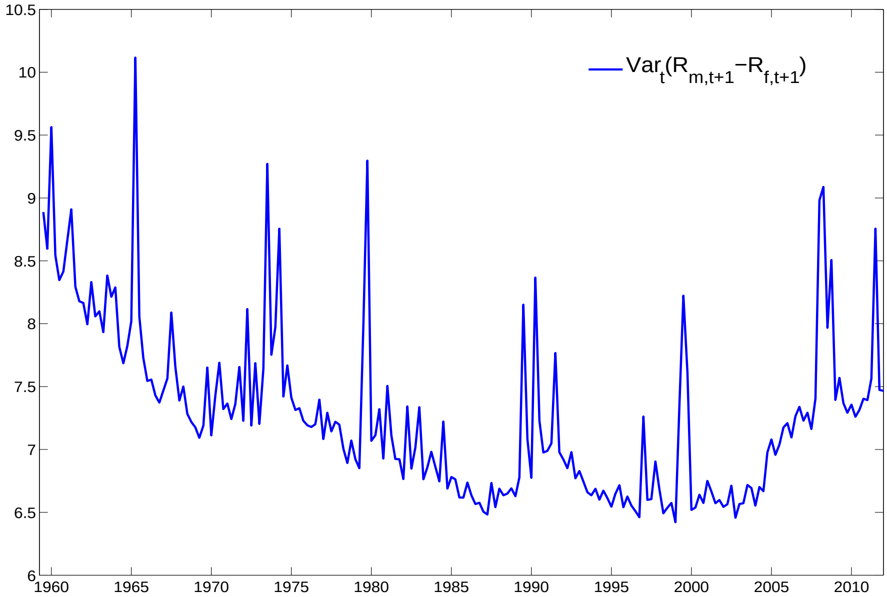


Figure C.8: The Expected State of Macroeconomic Risk

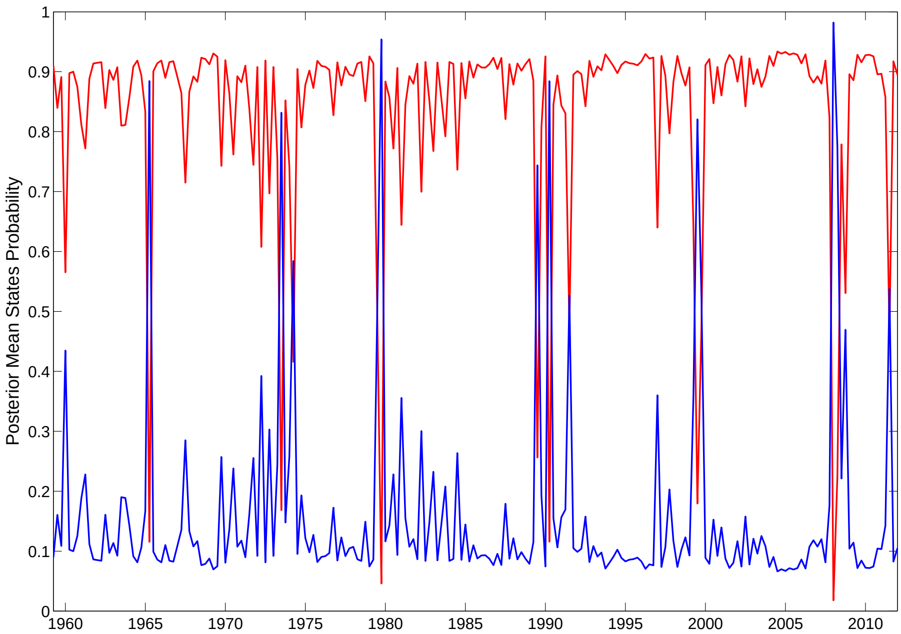


Figure C.9: Conditional Expected Growth Rate of Consumption

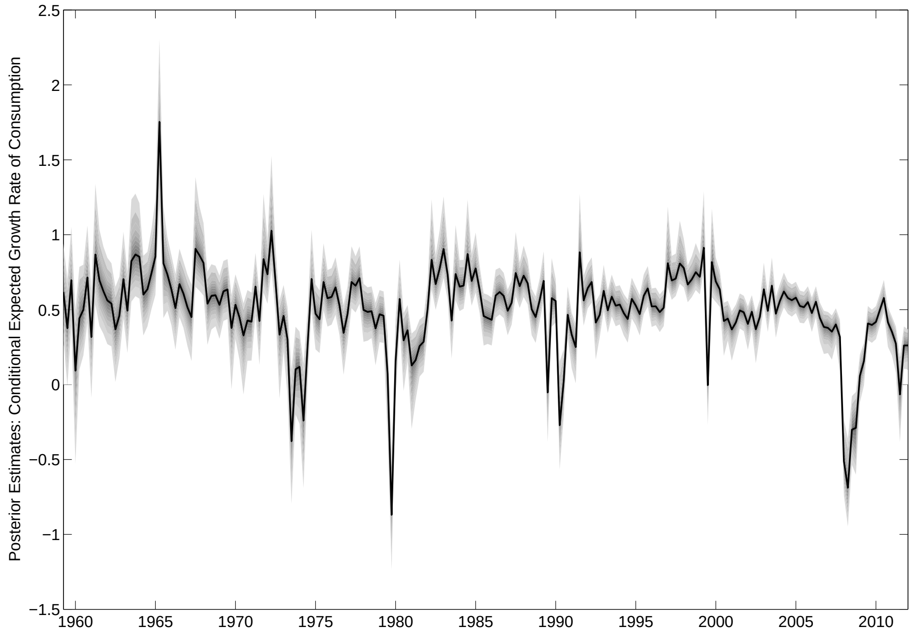


Figure C.10: Persistence

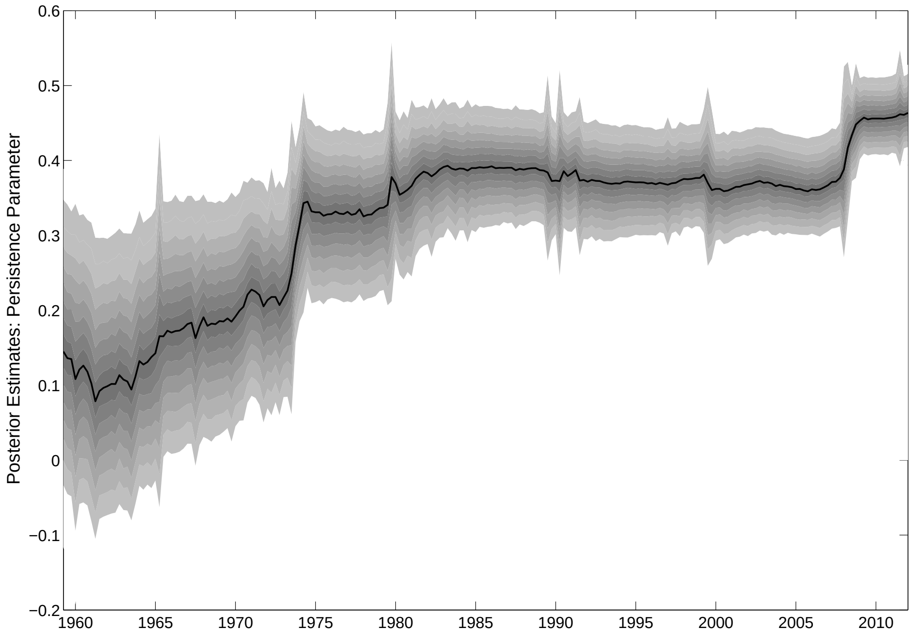


Figure C.11: Long-Run Expected Growth Rate of Consumption

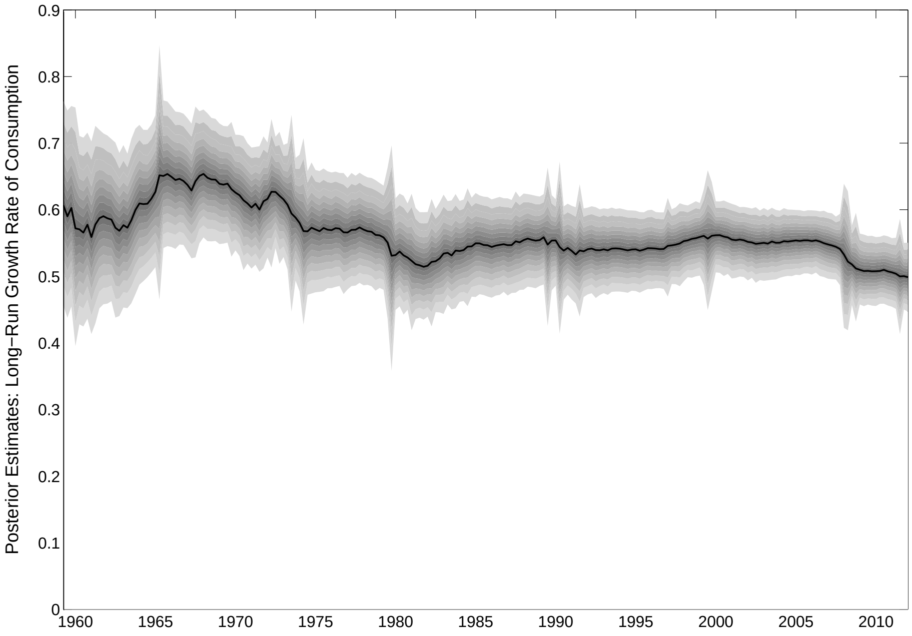


Figure C.12: Long-Run Expected Growth Rate of Dividends

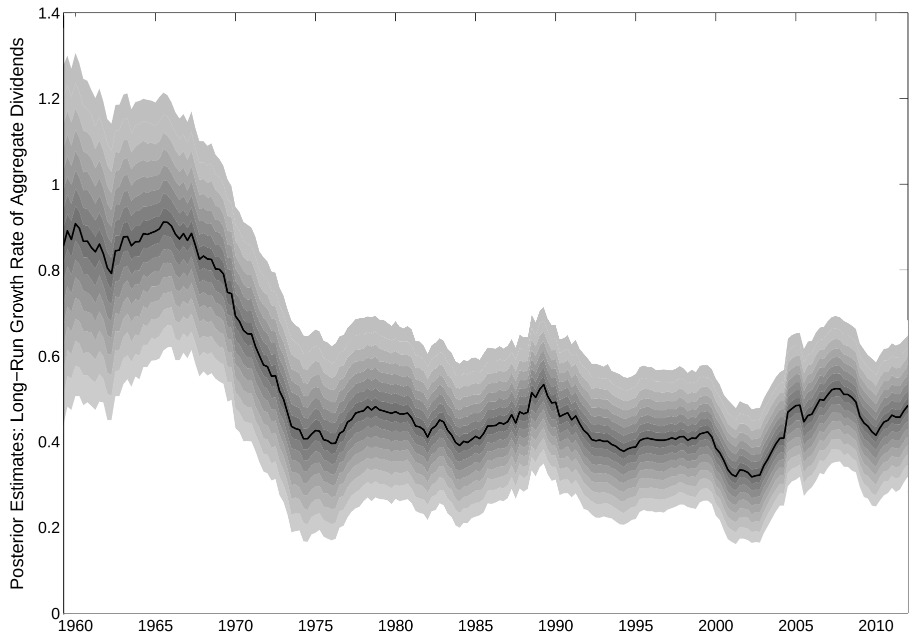
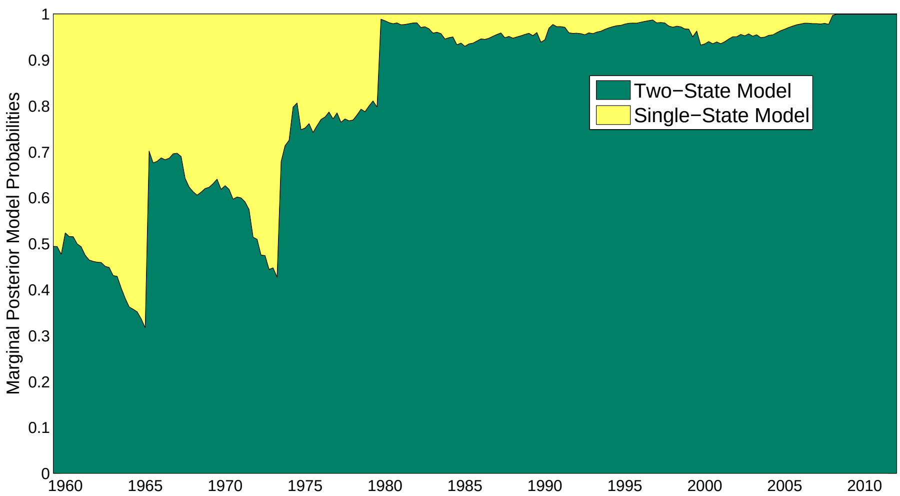


Figure C.13: Marginal Posterior Model Probabilities



Bibliography

Abel, A. 1999. Risk Premia and Term Premia in General Equilibrium. *Journal of Monetary Economics* 43:3–33.

Ang, A., and J. Chen. 2007. CAPM Over the Long-Run: 1926-2001. *Journal of Empirical Finance* 14:1–40.

Baker, S., N. Bloom, and S. Davis. 2013. Measuring Economic Policy Uncertainty. *Working Paper* .

Bakshi, G., and N. Kapadia. 2003. Delta-Hedged Gains and the Negative Market Volatility Risk Premium. *Review of Financial Studies* 16:527–566.

Bakshi, G., and D. Madan. 2006. A theory of Volatility Spreads. *Management Science* 52:1945–1956.

Bakshi, G., and G. Skoulakis. 2010. Do Subjective Expectations Explain Asset Pricing Puzzles? *Journal of Financial Economics* 98:462–477.

Bali, T., and A. Hovakimian. 2009. Volatility Spreads and Expected Stock Returns. *Management Science* 55:1797–1812.

Bansal, R., R. F. Dittmar, and C. T. Lundblad. 2005. Consumption, Dividends and the Cross-Section of Equity Returns. *The Journal of Finance* LX:1639–1672.

Bansal, R., D. Kiku, and A. Yaron. 2007. Risks for the Long-Run: Estimation and Inference. *Unpublished paper* .

Bansal, R., and I. Shaliastovich. 2011. Learning and Asset-Price Jumps. *The Review of Financial Studies* 8.

Bansal, R., and A. Yaron. 2004. Risk for the Long-Run: A Potential Resolution of Asset Pricing Puzzles. *The Journal of Finance* LIX:1481–1509.

Barndorff-Nielsen, O., and N. Shepard. 2002. Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society: Series B* 64:253–280.

Barsky, R., and J. De Long. 1993. Why Does Stock Market Fluctuates? *Quarterly Journal of Economics* 108:291–311.

Bekaert, G., E. Engstrom, and Y. Xing. 2009. Risk, Uncertainty, and Asset Prices. *Journal of Financial Economics* 91:59–82.

Bekaert, G., R. Hodrick, and X. Zhang. 2012. Aggregate Idiosyncratic Volatility. *Journal of Financial and Quantitative Analysis* 47:1155–1185.

Bekaert, G., and M. Hoerova. 2013. The VIX, the Variance Premium and Stock Market Volatility. *Netspar Discussion Paper* .

Benzoni, L., P. Collin-Dufresne, and R. Goldstein. 2011. Explaining Asset Pricing Puzzles Associated With the 1987 Market Crash. *Journal of Financial Economics* 101:552–573.

Bloom, N. 2009. The impact of uncertainty shocks. *Econometrica* 77:623–685.

Bollerslev, T., M. Gibson, and H. Zhou. 2011. Dynamics Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities. *Journal of Econometrics* 160:102–118.

Bollerslev, T., G. Tauchen, and H. Zhou. 2009. Expected Stock Returns and Variance Risk Premia. *Review of Financial Studies* 22:4463–4492.

Bollerslev, T., and V. Todorov. 2011. Tails, Fears and Risk Premia. *The Journal of Finance* 66:2165–2211.

Brandt, M., Q. Zeng, and L. Zhang. 2004. Equilibrium stock return dynamics under alternative rules of learning about hidden states. *Journal of Economics Dynamics and Control* pp. 1925–1954.

Breeden, D., M. Gibbons, and R. Litzenberger. 1989. Empirical Tests of the Consumption-Oriented CAPM. *Journal of Finance* 44:231–262.

Brennan, M., and Y. Xia. 2002. Dynamic Asset Allocation Under Inflation. *Journal of Finance* 57:1201–1238.

Britten-Jones, M., and A. Neuberger. 2000. Option Prices, Implied Price Processes, and Stochastic Volatility. *The Journal of Finance* LV:839–866.

Brodie, M., M. Chernov, and M. Johannes. 2007. Model Specification and Risk Premiums: Evidence From Future Options. *The Journal of Finance* pp. 839–866.

Burmeister, E., and M. McElroy. 1988. Joint Estimation of Factor Sensitivities and Risk Premia for the Arbitrage Pricing Theory. *Journal of Finance* 43:721–733.

Campbell, J. 1986. Bond and stock returns in a simple exchange model. *Quarterly Journal of Economics* pp. 785–804.

Campbell, J., and J. Beeler. 2012. The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment. *Critical Finance Review* 1:141–182.

Campbell, J., and J. Cochrane. 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy* 107:205–251.

Campbell, J., and R. Shiller. 1988a. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies* 1:195–227.

Campbell, J., and R. Shiller. 1988b. Stock Prices, Earnings, and Expected Dividends. *The Journal of Finance* 43:661–676.

Campbell, J., and L. Viceira. 1999. Consumption and portfolio decisions when returns are time-varying. *Quarterly Journal of Economics* pp. 433–495.

Campbell, L. M. M. B., J.Y., and Y. Xu. 2001. Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk. *Journal of Finance* 56:1–43.

Carceles-Poveda, E., and C. Giannitsarou. 2008. Asset Pricing with Adaptive Learning. *Review of Economic Dynamics* 11:629–651.

Carr, P., and L. Wu. 2009. Variance Risk Premiums. *Review of Financial Studies* 22:1311–1341.

Carter, C., and R. Kohn. 1994. On Gibbs Sampling for State-Space Models. *Biometrika* pp. 541–553.

Carvalho, C., M. Johannes, H. Lopes, and N. Polson. 2010a. Particle Learning and Smoothing. *Statistical Science* 25:88–106.

Carvalho, C., M. Johannes, H. Lopes, and N. Polson. 2010b. Particle Learning for Sequential Bayesian Computation. *Bayesian Statistics* 9.

Cecchetti, S., P. Lam, and N. Mark. 1999. Asset Pricing with Distorted Beliefs: Are Equity Returns "Too Good to be True"? *American Economic Review* pp. 787–805.

Chan, C. N., L.K.C., and D. Hsieh. 1985. An Exploratory Investigation of the Firm Size Effect. *Journal of Financial Economics* pp. 451–471.

Chan, K. J., L.K.C., and J. Lakonishok. 1998. The Risk and Return from Factors. *Journal of Financial and Quantitative Analysis* pp. 159–188.

Chan, W., and J. Maheu. 2002. Conditional Jump Dynamics in Stock Market Returns. *Journal of Business and Economic Statistics* 20:377–389.

Chen, H., W. Dou, and L. Kogan. 2013. Measuring the "Dark Matter" in Asset Pricing Models. *Working Paper* .

Chen, R. R., N.-F., and S. Ross. 1986. Economic Forces and the Stock Market. *Journal of Business* 59:383–403.

Chib, S. 1995. Marginal Likelihood from the Gibbs Output. *Journal of the American Statistical Association* 432:1313–1321.

Cochrane, J. 1999. Portfolio advice for a multifactor world. *Economic Perspectives Federal Reserve Bank of Chicago* 3:59–78.

Cochrane, J. 2005. *Asset Pricing*. Princeton, NJ: Princeton University Press.

Cochrane, J., and M. Piazzesi. 2005. Bond Risk Premia. *American Economic Review* 94:138–160.

Cogley, T., and T. Sargent. 2008. The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression? *Journal of Monetary Economics* 55:454–476.

Cogley, T., and T. Sargent. 2009. Anticipated Utility and Rational Expectations as Approximations of Bayesian Decision Making. *International Economic Review* 49:185–221.

Collin-Dufresne, C., M. Johannes, and L. Lochstoer. 2013. Parameter Learning in General Equilibrium: The Asset Pricing Implications. *Columbia Business School, Working paper series* .

Connor, G., and R. Korajczyk. 1998. Risk and Return in an Equilibrium APT: Application of a New Test Methodology. *Journal of Financial Economics* 21:255–289.

Corradi, V., W. Distaso, and A. Mele. 2013. Macroeconomic Determinants of Stock Volatility and Volatility Premiums. *Journal of Monetary Economics* 60:203–220.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou. 1999. A Guide to Volatility and Variance Swaps. *Journal of Derivatives* 4:9–32.

Drechsler, I. 2013. Uncertainty, Time-Varying Fear, and Asset Prices. *The Journal of Finance* 68:1843–1889.

Drechsler, I., and A. Yaron. 2011. What's Vol Got to do With It. *Review of Financial Studies* 24:1–45.

Epstein, L., and S. Zin. 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57:937–969.

Epstein, L., and S. Zin. 1991. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation. *Journal of Political Economy* 99:263–286.

Fama, E., and K. French. 1988a. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* pp. 23–49.

Fama, E., and K. French. 1988b. Dividend Yields and Expected Stock Returns. *Journal of Financial Economics* 22:3–25.

Fama, E., and J. MacBeth. 1973. Risk, Return and Equilibrium: Empirical Tests. *Journal of Political Economy* 81:607–636.

Ferson, W., and C. Harvey. 1991. The Variation of Economic Risk Premiums. *Journal of Political Economy* 99:385–415.

Ferson, W., and R. Korajczyk. 1995. Do Arbitrage Pricing Models Explain the Predictability of Stock Returns? *Journal of Business* 68:309–349.

Flannery, M., and A. Protopapadakis. 2002. Macroeconomic Factors Do Influence Aggregate Stock Returns. *Review of Financial Studies* 15:751–782.

Frühwirth-Schnatter, S. 1994. Data Augmentation and Dynamic Linear Models. *Journal of Time Series Analysis* 15:183–202.

Geman, S., and D. Geman. 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions* 6:721–741.

Gerlach, C. C., R., and R. Kohn. 2000. Efficient Bayesian Inference for Dynamic Mixture Models. *Journal of the American Statistical Association* 95:819–828.

Geweke, J. 2001. A Note on Some Limitations of CRRA Utility. *Economics Letters* 71:341–345.

Geweke, J., and G. Zhou. 1996. Measuring the Pricing Error of the Arbitrage Pricing Theory. *Review of Financial Studies* 9:557–587.

Ghysels, E. 1998. On Stable Factor Structure in the Pricing: Do Time-Varying Betas Help or Hurt? *Journal of Finance* 53:549–573.

Giordani, P., and R. Kohn. 2008. Efficient Bayesian Inference for Multiple Change-Point and Mixture Innovation Models. *Journal of Business and Economic Statistics* 26:66–77.

Guidolin, M., and A. Timmermann. 2007. Properties of Equilibrium Asset Prices Under Alternative Learning Schemes. *Journal of Economic Dynamics and Control* 31:161–217.

Hamilton, J. 1994. *Time series analysis*. Princeton University Press.

Han, B., and Y. Zhou. 2013. Variance Risk Premium and Cross-Section of Stock Returns. *Working Paper* .

Hansen, L. 2007. Beliefs, Doubts and Learning: Valuing Macroeconomic Risk. *The American Economic Review* 97:1–30.

Hansen, L., J. Heaton, and N. Li. 2008. Consumption Strikes Back? Measuring Long-Run Risk. *Journal of Political Economy* 116:260–302.

Hansen, L., and T. Sargent. 2010. Fragile Beliefs and the Price of Uncertainty. *Quantitative Economics* 1:129–162.

Harvey, A. 1981. *Time series models*. Philip Allan, London.

Harvey, A. 1989. *Forecasting, structural time series models and the Kalman Filter*. Cambridge University Press.

Harvey, C. R. 2001. The Specification of Conditional Expectations. *Journal of Empirical Finance* 8:573–638.

Henkel, M. J., S.J., and F. Nardari. 2011. Time-Varying Short-Horizon Predictability. *Journal of Financial Economis* 99:560–580.

Jagannathan, R., and Z. Wang. 1998. An Asymptotic Theory for Estimating Beta-Pricing Models Using Cross-Sectional Regression. *Journal of Finance* 53:1285–1309.

Johannes, M., A. Korteweg, and N. Polson. 2013. Sequential Learning, Predictive Regressions, and Optimal Portfolio Returns. *The Journal of Finance* forthcoming.

Johannes, M., L. Lochstoer, and Y. Mou. 2011. Learning About Consumption Dynamics. *Columbia Business School, Working Paper* .

Johannes, M., and N. Polson. 2005. MCMC Methods for Financial Econometrics. *Handbook of Financial Econometrics* .

Jostova, G., and A. Philipov. 2005. Bayesian Analysis of Stochastic Betas. *Journal of Financial and Quantitative Analysis* 40:747–778.

Ju, N., and J. Miao. 2012. Ambiguity, Learning, and Asset Returns. *Econometrica* 80:559–591.

Kass, R., and A. Raftery. 1995. Bayes Factors. *Journal of the American Statistical Association* 430:773–795.

Koop, G., and S. Potter. 2007. Estimation and Forecasting with Multiple Breaks. *Review of Economic Studies* 74:763–789.

Kramer, C. 1994. Macroeconomic Seasonality and the January Effect. *Journal of Finance* 49:1883–1891.

Kreps, D. 1998. Anticipated Utility and Dynamic Choice. *in D.P.Jacobs, E.Kalai, and M.Kamien, eds., Frontiers of Research in Economic Theory (Cambridge university press)* pp. 242–274.

Kreps, D., and E. Porteus. 1978. Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica* 46:185–200.

Lamont, O. 2001. Economic Tracking Portfolios. *Journal of Finance* 105:161–184.

Lettau, M., S. Ludvigson, and J. Watcher. 2008. The Declining Equity Premium: What Role Does Macroeconomic Risk Play? *Review of Financial Studies* 21:1653–1687.

Lewellen, J., and J. Shanken. 2002. Learning, Asset-Pricing Tests, and Market Efficiency. *The Journal of Finance* 57:1113–1145.

Liu, J., J. Pan, and T. Wang. 2005. An Equilibrium Model of Rare-Event Premia and its Implications for Option Smirks. *Review of Financial Studies* 18:131–164.

Ludvigson, S., K. Jurado, and S. Ng. 2013. Measuring Uncertainty. *Working Paper* .

Ludvigson, S., and S. Ng. 2009. Macro Factors in Bond Risk Premia. *Review of Financial Studies* 22:5027–5067.

Maheu, J., and T. McCurdy. 2009. How Useful are Historical Data for Forecasting the Long-Run Equity Return Distribution? *Journal of Business and Economic Statistics* 27:95–112.

Maheu, J. M., and S. Gordon. 2008. Learning, Forecasting and Structural Breaks. *Journal of Applied Econometrics* 23:553–583.

Martin, I. 2013. Consumption-Based Asset Pricing With Higher Cumulants. *Review of Economic Studies* 80:745–773.

Massa, M., and A. Simonov. 2005. Is learning a Dimension of Risk? *Journal of Banking and Finance* 29:2605–2632.

McCulloch, R., and P. Rossi. 1990. Posterior, Predictive and Utility Based Approaches to Testing Arbitrage Pricing Theory. *Journal of Financial Economics* 28:7–38.

McQueen, G., and V. Roley. 1993. Stock Prices, News, and Business Conditions. *Review of Financial Studies* 6:683–707.

Meddahi, N. 2002. Theoretical Comparison Between Integrated and Realized Volatility. *Journal of the Applied Econometrics* 17:479–508.

Mehra, R., and E. Prescott. 1985. The Equity Premium: A Puzzle. *Journal of Monetary Economics* 15:146–161.

Mei, J. 1993. A Semiautoregression Approach to the Arbitrage Pricing Theory. *Journal of Finance* 48:599–620.

Miao, J., W. Bin, and H. Zhou. 2012. Ambiguity Aversion and Variance Premium. *Working Paper* .

Miazhynskaia, F.-S. S. D. G., T. 1988. Bayesian Variable Selection in Linear Regression. *Journal of the American Statistical Association* 404:1023–1032.

Miazhynskaia, F.-S. S. D. G., T. 2006. Bayesian Testing for Non-Linearity in Volatility Modeling. *Computational Statistics & Data Analysis* 51:2029–2042.

Moore, B., and H. Schaller. 1996. Learning, regime switches, and equilibrium asset pricing dynamics. *Journal of Economic Dynamics and Control* pp. 979–1006.

Næ s, S. J., R., and B. Ø degaard. 2011. Stock Market Liquidity and the Business Cycle. *Journal of Finance* 66:139–176.

Nardari, F., and J. Scruggs. 2007. Bayesian Analysis of Linear Factor Models with Latent Factors, Multivariate Stochastic Volatility, and APT Pricing Restrictions. *Journal of Financial and Quantitative Analysis* 42:857–891.

Pagan, A. 1984. Econometric Issues in the Analysis of Regressions with Generated Regressors. *International Economic Review* 25:221–247.

Pastor, L., and F. Stambaugh, R. 2003. Liquidity Risk and Expected Stock Returns. *Journal of Political Economy* 111:642–685.

Piazzesi, S., and M. Schneider. 2010. Trend and Cycle in Bond Risk Premia. *Working paper series, Stanford University* .

Sargent, T. 2007. Commentary: Long-Run Risks and Financial Markets 4.

Shanken, J. 1992. On the Estimation of Beta Pricing Models. *Review of Financial Studies* 5:1–34.

Shanken, J., and M. Weinstein. 2006. Economic Forces and the Stock Market Revisited. *Journal of Empirical Finance* 13:129–144.

Shiller, R. 1981. Do Stock Prices Move Too Much to be Justified by Subsequent Change in Dividends? *American Economic Review* 71:421–436.

Tauchen, G. 2012. Stochastic Volatility in General Equilibrium. *Quarterly journal of finance* 1.

Timmermann, A. 1993. How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices. *Quarterly Journal of Economics* 108:1135–1145.

Timmermann, A. 1996. Excess Volatility and Predictability of Stock Prices in Autoregressive Dividend Models with Learning. *Review of Economic Studies* 63:523–557.

Todorov, V. 2010. Variance Risk-Premium Dynamics: The Role of Jumps. *Review of Financial Studies* 23:345–383.

Tu, J. 2010. Is Regime Switching in Stock Returns Important in Portfolio Decisions? *Management Science* 56:1198–1215.

Valkanov, R. 2003. Long-Horizon Regressions: Theoretical Results and Applications. *Journal of Financial Economics* 68:201–232.

Veronesi, P. 1999. Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model. *Review of Financial Studies* 12:975–1007.

Veronesi, P. 2000. How Does Information Quality Affect Stock Returns? *The Journal of Finance* 2:807–837.

Weil, P. 1989. The Equity Premium Puzzle and the Risk-Free Rate Puzzle. *Journal of Monetary Economics* 24:401–421.

Weitzman, M. 2007. Subjective Expectations and Asset-Return Puzzles. *American Economic Review* 97:1102–1130.

West, M., and J. Harrison. 1997. *Bayesian forecasting and dynamics models*. Springer.

Williams, N. 2003. Adaptive Learning and Business Cycles. *Working paper, Princeton* .

Wu, L., and F. Zhang. 2008. A No-Arbitrage Analysis of Macroeconomic Determinants of the Credit Spread Term Structure. *Management Science* 54:1160–1175.

Zhou, G. 1999. Security Factors as Linear Combinations of Economic Variables. *Journal of Financial Markets* 2:403–432.