

Robust Equilibrium Yield Curves

Isaac KLESHCHELSKI
Nicolas VINCENT

Cahier de recherche n° IEA-08-02

November 2007

Copyright © 2008 HEC Montréal.

Tous droits réservés pour tous pays. Toute traduction ou toute reproduction sous quelque forme que ce soit est interdite.

Les textes publiés dans la série des Cahiers de recherche HEC n'engagent que la responsabilité de leurs auteurs.

La publication de ce Cahier de recherche a été rendue possible grâce à des subventions d'aide à la publication et à la diffusion de la recherche provenant des fonds de l'École des HEC.

Direction de la recherche, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal (Québec) Canada H3T 2A7.

Robust Equilibrium Yield Curves*

Isaac Kleshchelski,[†] Nicolas Vincent[‡]

November 2007

Abstract

This paper studies the quantitative implications of the interaction between robust control and stochastic volatility for key asset pricing phenomena. We present an equilibrium term structure model with a representative agent and an output growth process that is conditionally heteroskedastic. The agent does not know the true model of the economy and chooses optimal policies that are robust to model misspecification. The choice of robust policies greatly amplifies the effect of conditional heteroskedasticity in consumption growth, improving the model's ability to explain asset prices. In a robust control framework, stochastic volatility in consumption growth generates both a state-dependent market price of model uncertainty and a stochastic market price of risk. We estimate the model using data from the bond and equity markets, as well as consumption data. We show that the model is consistent with key empirical regularities that characterize the bond and equity markets. We also characterize empirically the set of models the robust representative agent entertains, and show that this set is 'small'. That is, it is statistically difficult to distinguish between models in this set.

JEL classification: D81, E43, G11, G12

Keywords: Yield curves, Market price of Uncertainty, Robust control

*We are especially indebted to Martin Eichenbaum and Sergio Rebelo for their continuous advice, support and encouragement. We also thank Torben Andersen, Craig Burnside, Lars Hansen, Zhiguo He, Ravi Jagannathan, Arvind Krishnamurthy, Jonathan Parker, Giorgio Primiceri, Costis Skiadas, Annette Vissing-Jorgensen, Ernst Schaumburg and participants in the Macro Seminar and Finance Bag Lunch at Northwestern University.

[†]Kellogg School of Management, Northwestern University, 2001 Sheridan Rd., Evanston, IL 60208-2001.
Email: ikleshchelski@kellogg.northwestern.edu

[‡]HEC Montréal and CIRPÉE, 3000 chemin de la Côte-Sainte-Catherine, Montréal, QC, Canada H3T 2A7.
Email: nicolas.vincent@hec.ca

1 Introduction

This paper studies the implications of the interaction between robust control and stochastic volatility for key asset pricing phenomena. We quantitatively show that robustness, or fear of model misspecification, coupled with state-dependent volatility provides an empirically plausible characterization of the level and volatility of the equity premium, the risk free rate, and the cross-section of yields on treasury bonds. We also show that robustness offers a novel way of reconciling the shape of the term structure of interest rates with the persistence of yields. Finally, we quantify the level of robustness encoded in the agents' behavior.

We construct a continuous-time, Lucas tree, asset pricing model in which a representative agent is averse to both risk and ambiguity. The presence of ambiguity stems from the agent's incomplete information about the economy's data generating process (DGP). In other words, the agent does not know which of several possible models is the true representation of the economy. Introducing ambiguity aversion into our framework allows us to reinterpret an important fraction of the market price of risk as the market price of model (or Knightian) uncertainty. We model ambiguity aversion using robust control techniques as in Anderson et al. (2003).¹ In our model, the representative agent distrusts the reference model and optimally chooses a distorted DGP. His consumption and portfolio decisions are then based on this distorted distribution. Ambiguity aversion gives rise to endogenous pessimistic assessments of the future.

A key assumption in the model is that the output growth process is conditionally heteroskedastic. The consumption growth process inherits this heteroskedasticity, which gives rise to a stochastic market price of risk.² The main contribution of this paper is to show that ambiguity aversion greatly amplifies the effect of stochastic volatility in consumption growth and, therefore, can explain asset prices in an empirically plausible way. In the absence of ambiguity aversion, plausible degrees of stochastic volatility in consumption growth do not generate sufficient variation in the stochastic discount factor.

By choosing a distorted DGP, the robust representative agent has biased expectations of future consumption growth. Being pessimistic, the agent tilts his subjective distribution towards

¹Behavioral puzzles such as the Ellsberg paradox (Ellsberg (1961)) led to the axiomatization of the maxmin decision making by Gilboa and Schmeidler (1989). Robust control is one way of modeling Knightian uncertainty. For a comprehensive treatment of robustness see Hansen and Sargent (2007a). Examples of the use of robust control in economics and finance include Anderson et al. (2003), Cagetti et al. (2002), Gagliardini et al. (2004), Hansen and Sargent (2007b), Hansen et al. (2006), Liu et al. (2005), Maenhout (2004), Routledge and Zin (2001), Uppal and Wang (2003). An alternative approach to modeling ambiguity allows agents to have multiple priors. See, for example, Epstein and Schneider (2003), Epstein and Wang (1994).

²Recently, several authors (e.g., Bansal and Yaron (2004), Bansal et al. (2005)) argue that conditionally heteroskedastic consumption growth is potentially important to understand asset prices. We agree that this channel is indeed important, but claim that it is the interaction with ambiguity aversion that is critical.

states in which marginal utility is high. With stochastic volatility, positive volatility surges result in a more diffuse distribution of future consumption growth. In that case, the objective distribution assigns more probability mass to future ‘bad’ realizations of consumption growth. The agent seeks policies that can reasonably guard against such ‘bad’ realizations. Consequently, he increases the distortion to his expectations of consumption growth. The interaction between robustness and stochastic volatility introduces a state dependent distortion to the drift of consumption growth, and therefore, to the drift in the agent’s intertemporal marginal rate of substitution.³ This state dependent distortion generates sharp implications for asset pricing phenomena.

We estimate our model and assess its implications using data from the equity and bond markets, as well as consumption data. We exploit cross-equation restrictions across bond and equity markets to improve both the identification of structural parameters in our model and the estimation of the market price of risk and uncertainty.⁴ Our main findings are as follows.

First, we show that our model, calibrated with a unitary degree of risk aversion and elasticity of intertemporal substitution (EIS), can reproduce both the high and volatile equity premium and the low and stable risk free rate observed in the data. Previous studies, such as Mehra and Prescott (1985) and Weil (1989), have shown that explaining the behavior of the equity premium requires implausibly high levels of risk aversion. Ambiguity aversion generates an uncertainty premium that helps to alleviate the difficulties these previous studies have encountered. Since there is no benchmark value for the degree of ambiguity aversion, we use detection error probabilities to show that the degree of robustness required to fit the data is reasonable. In other words, we empirically show that the set of models the robust representative agent entertains is small. That is, it is statistically difficult to distinguish between models in this set.

Second, our model can account for the means of the cross-section of bond yields. In particular, we can replicate the upward sloping unconditional yield curve observed in the data. This result highlights a novel interpretation of the uncertainty premium generated by robustness. On empirical grounds, we assume that the conditional variance of output growth, and hence con-

³In the next section and in the empirical section of the paper we provide an additional, more technical, explanation for the link between stochastic volatility and the robust distortion by using the properties of the relative entropy and the size of the set of models the agent entertains.

⁴Campbell (2000) convincingly argues that "it is important to reconcile the characterization of the SDF provided by bond market data with the evidence from stock market data. Term structure models of the SDF are ultimately unsatisfactory unless they can be related to the deeper general-equilibrium structure of the economy. Researchers often calibrate equilibrium models to fit stock market data alone, but this is a mistake because bonds carry equally useful information about the SDF. The short-term real interest rate is closely related to the conditional expected SDF and thus to the expected growth rate of marginal utility; in a representative-agent model with power utility of consumption, this links the real interest rate to expected consumption growth...The risk premium on long-term bonds is also informative."

sumption growth, is stationary and positively correlated with the consumption growth process. This positive correlation implies that when marginal utility is high the conditional variance of consumption growth is low. Consequently, a downward bias in the subjective conditional expectations of consumption growth induces a negative distortion to the subjective expectations of variance changes. We show that this negative distortion is a linear function of the level of the conditional variance process. Consequently, the unconditional distortion is a linear negative function of the objective steady state of the variance process. Therefore, the subjective steady state of the variance process is lower than the objective steady state. In other words, on average, the agent thinks that the conditional variance of consumption growth should decrease. Since the unconditional level of bond yields and the steady state level of the conditional volatility of consumption growth are inversely related, the agent expects, on average, that yields will increase. Consequently, the unconditional yield curve is upward sloping.

Third, our model can replicate the declining term structure of unconditional volatilities of real yields, and the negative correlation between the level and the spread of the real yield curve. The fact that the robust distortion to the conditional variance process is a linear function of the level of the variance implies that the distorted process retains the mean reversion structure of the objective process. Since shocks to the conditional variance are transitory, the short end of the yield curve is more responsive to volatility shocks relative to the long end. Hence, short maturity yields are more volatile than long maturity yields. Also, our model implies that yields are an affine function of the conditional variance of consumption growth. Therefore, all yields are perfectly positively correlated. Since short yields are more responsive to volatility shocks relative to long yields, but both move in the same direction, when yields decrease the spread between long yields and short yields increases and becomes more positive. Thus, the level and spread of the real yield curve are negatively correlated.

Fourth, the model can reconcile two seemingly contradictory bond market regularities: the strong concavity of the short end of the yield curve and the high degree of serial correlation in bond yields.⁵ The intuition for this result is closely linked to the mechanism behind the upward sloping real yield curve. Generally, in a one-factor affine term structure model, the serial correlation of yields is driven by the serial correlation of the state variable implied by the objective DGP. In contrast, in our model the slope of the yield curve is shaped by the degree of mean reversion of the conditional variance process implied by the agent's distorted (i.e., subjective) distribution. The state dependent distortion to the variance process not only changes the perceived steady state of the variance but also its velocity of reversion. With

⁵In a standard one-factor model, it is difficult to separate these two properties, since both observations are directly tied to the persistence of the underlying univariate shock process.

positive correlation between consumption growth and the conditional variance process, we show that the subjective mean reversion is faster than the objective one. Ex ante, the agent expects shocks to the variance process to die out fast, but ex-post these shocks have a longer lasting effect than expected. The slope of the yield curve is a reflection of how fast the agent expects the effect of variance surges to dissipate. Hence, the positive slope of the yield curve is rapidly declining when the subjective mean reversion is high. However, the serial correlation of yields is measured ex-post. Hence, if the objective persistence of the variance process is high, yields are highly persistent, which is in line with the empirical evidence. When the agent seeks more robustness, the separation between the ex-ante and ex-post persistence is stronger.

The remainder of the paper is organized as follows. In section 2 we present the robust control idea in a simple two-period asset pricing model. The main purpose of this section is to highlight the channels through which uncertainty aversion considerations alter the predictions of a standard asset pricing model. In section 3 we present our continuous time model and discuss its implications for the equity market's valuation patterns and the implied risk free rate. In section 4 we discuss the model's predictions concerning the bond market. We derive analytical affine term structure pricing rules and discuss the distinction between the market price of risk and uncertainty. In section 5 we present empirical evidence that supports our modeling assumptions. We also estimate our complete model and investigate the implied level of uncertainty aversion exhibited by the representative agent. In section 6 we offer our concluding remarks and discuss potential avenues for future research.

2 Robustness in a Two-Period Example

In this section we introduce the terminology and concepts used throughout the paper using a two-period consumption-saving example. This example helps build our intuition and motivate the modeling assumptions used in our model.

2.1 Reference and Distorted Models

The representative agent in our economy uses a *reference* or *approximating* model. However, since he fears that this model is potentially misspecified, he chooses to diverge from it when making his decisions.⁶ In the context of this paper, the reference model is assumed to generate

⁶Another possibility is to claim that for some reason the agent dislikes extreme negative events and wants to take special precautionary measures against these events. If we choose this behavioral interpretation, we can then assume the agent knows the true DGP, but that his marginal utility function is very high in bad states of the world. Low consumption is so costly that the agent requires policies that are robust to these states. Even though there is complete observational equivalence between the two approaches, they are utterly different from

the observed data. In contrast with the rational expectations paradigm, the agent entertains alternative DGPs. The size of the set of possible models is implicitly defined by a penalty function (relative entropy) incorporated into the agent's utility function. So the agent chooses an optimal distribution for the exogenous processes. In other words, the agent optimally chooses his set of beliefs simultaneously with the usual consumption and investment decisions. The robust agent distorts the approximating model in a way that allows him to incorporate fear of model misspecification. We will refer to the optimally chosen model as the *distorted* model.⁷

2.2 A Two-Period Model

We now discuss a simple two-period example. Our discussion is intentionally informal. Our goal is to illustrate how robustness considerations alter the predictions of a standard asset pricing model. We consider a Lucas-tree type economy in which the agent receives one unit of consumption good in the first period. He decides how many units (α) of a claim to the stochastic endowment in period one (D_1) to buy. The unit price of a claim to the tree's output is denoted by S_0 . We assume that period one output is drawn from a lognormal distribution, which is referred to as \mathbb{P} :

$$\ln D_1 \sim N(\mu, \sigma^2).$$

We denote the agent's subjective distribution by \mathbb{Q} . The robust agent solves a max-min problem, where the minimization takes place over \mathbb{Q} :⁸

$$\begin{aligned} & \max_{\alpha} \min_{\mathbb{Q}} \{u(C_0) + \beta \mathbb{E}^{\mathbb{Q}}[u(C_1) + \theta \mathcal{R}(\mathbb{Q})]\} \\ & \text{subject to:} \\ & C_0 = 1 - \alpha S_0; \quad C_1 = \alpha D_1. \end{aligned} \tag{2.1}$$

Here C_0 and C_1 are the levels of consumption in periods zero and one, respectively. The object $\mathcal{R}(\mathbb{Q})$ represents the penalty imposed on the agent whenever he decides to choose a distribution different from \mathbb{P} . We assume that this penalty is the relative entropy, or Kullback-Leibler divergence, between the objective (\mathbb{P}) and subjective (\mathbb{Q}) distributions. The parameter

a behavioral perspective.

⁷An alternative is to allow for the possibility that a different, unspecified model, is actually the DGP. In this scenario, it is likely that neither the distorted nor the reference model generate the data. The agent must in this case infer which model is more likely to generate the data. See Hansen and Sargent (2007b) for an example.

⁸This preference specification is referred to in the literature as 'multiplier preferences'. The decision theoretic foundation for the use of multiplier preferences is discussed in Maccheroni et al. (2006) and Strzalecki (2007). These authors also discuss the interpretation of the parameter θ as a measure of the level of ambiguity aversion which the agent exhibits.

θ is a multiplier which determines the sensitivity of the agent's value function to the relative entropy. Without this penalty, the minimization problem would have a boundary solution in which the agent assigns all the probability mass to the worst possible state (if the support is the entire real line, the agent will distort the mean to negative infinity). Note also that when $\theta \rightarrow \infty$, problem (2.1) converges to the conventional time-additive expected utility case. In this case, the penalty for distorting the objective distribution is so large, that the agent optimally decides to construct his beliefs using the objective measure \mathbb{P} (since $\mathcal{R}(\mathbb{Q}) = 0$ when $\mathbb{P} = \mathbb{Q}$).

In continuous time, one can show that given a conditional normal distribution for the growth rate of output, the distorted distribution is also normal with the same variance and a lower mean. This result follows from the regularity conditions (i.e., absolute continuity) required when using relative entropy.⁹ Since our complete model is cast in continuous time, we consider only mean distortions in this two-period example in order to make the transition to the full model more transparent. The following lemma characterizes more generally what distortions the agent considers in discrete time over the class of univariate Normal distributions:

Lemma 1 *Consider the class of Normal distributions. If $u(C) = \ln(C)$ then the agent chooses $\mu_{\mathbb{Q}} = \mu + h$ and $\sigma_{\mathbb{Q}}^2 = \sigma^2$. If $u(C) = C^{1-\gamma}/(1-\gamma)$, $\gamma \neq 1$ then the agent chooses $\mu_{\mathbb{Q}} = f(\mu, \sigma^2)$ and $\sigma_{\mathbb{Q}}^2 = g(\mu, \sigma^2)$ for some $f, g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ such that if $\mu > \mu_{\mathbb{Q}}$ then $\frac{\partial \sigma_{\mathbb{Q}}^2}{\partial (\mu - \mu_{\mathbb{Q}})} > 0$*

Proof. See Appendix A. ■

We assume that utility is logarithmic. Lemma 1 shows that, in this case, it is optimal for a robust agent to distort only the mean of the distribution. An agent with logarithmic utility derives no benefit from distorting the variance, since he cares only about the first moment of the distribution. But, he incurs a cost since an increase in variance raises the relative entropy of the two distributions. In contrast, in the non-logarithmic case ($\gamma \neq 1$), the agent also cares about the second moment of the distribution and therefore distorts the variance. We see that $\partial \sigma_{\mathbb{Q}}^2 / \partial (\mu - \mu_{\mathbb{Q}}) > 0$ since the ‘distance’ between the distributions is positively related to mean distortions. However, for a given mean distortion, the agent needs to distort the variance in order to maintain a desired distance between the distributions.

Let \mathbb{Q} be $N(\mu + h, \sigma^2)$, where $h \in \mathbb{R}$ represents the mean distortion chosen by the robust

⁹Relative entropy and Radon-Nikodym are ultimately likelihood ratios. In continuous time, likelihood ratios are extremely sensitive to variance distortions. This is because high frequency observations allow us to estimate the variance fairly accurately. So, the likelihood ratio reveals the difference between models with different variances easily. In the robust control formulation, this imposes a large penalty on the agent. Thus, he optimally chooses not to distort the variance.

agent. The relative entropy of \mathbb{P} and \mathbb{Q} is given by,

$$\begin{aligned}\mathcal{R}(\mathbb{Q}) &\equiv \int \ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) d\mathbb{Q}, \\ &= \frac{h^2}{2\sigma^2}.\end{aligned}\tag{2.2}$$

Not surprisingly, the divergence between \mathbb{P} and \mathbb{Q} is a positive function of the distortion to the mean, h . The presence of the variance in the denominator reflects the fact that the distortion in the mean must be measured relative to the degree of volatility associated with the distribution.

Assuming that $u(C) = \ln C$ and using (2.2) we can rewrite (2.1) as:

$$\max_{\alpha} \min_h \left\{ \ln(1 - \alpha S_0) + \beta \mathbb{E}^{\mathbb{Q}} \left[\ln(\alpha D_1) + \frac{\theta h^2}{2\sigma^2} \right] \right\}.$$

Note that now the minimization problem is taken over h which serves as a sufficient statistic for the divergence between \mathbb{P} and \mathbb{Q} . The first-order condition with respect to h yields:

$$h = -\frac{\sigma^2}{\theta} \leq 0.$$

This is the distortion to the mean of the conditional distribution of next period's output. As the penalty parameter θ becomes smaller, the agent seeks more robust policies and the (absolute) size of the distortion increases. A result that will prove particularly important in our context is that the distortion becomes more pronounced as the output is expected to be more volatile. The intuition for this result is rather simple: a robust agent is more prone to take precautionary measures against misspecification when bad outcomes are more likely.

Since h is independent of the other controls (i.e., neither consumption nor the investment policy affect the choice of h), the maximization with respect to α and the imposition of the equilibrium condition ($\alpha = 1$) yield the usual pricing formula:

$$S_0 = \frac{\beta}{1 + \beta}.$$

The maximization also yields the usual decision rule for consumption. The agent consumes a fraction of his wealth that is independent of the price of the risky asset:

$$C_0 = \frac{1}{1 + \beta}.$$

Note that up to this point, the robust considerations do not alter any of the model's predictions: both consumption and the price of the risky asset are unaffected by the choice of h . To understand this result, note that the optimal level of investment in the risky asset is affected by two forces. First, the robust agent fears exposing his capital to adverse shocks to the return on the risky asset. Second, the agent fears a bad output growth realization that will leave him 'hungry' next period and prefers to save. When momentary utility is logarithmic, these two effects cancel out.¹⁰

The implications of robustness in our setup take the form of a precautionary savings motive. To see this, first consider the Euler equation used to price a risk free one-period bond which is in equilibrium in zero net supply

$$1 = \beta \mathbb{E}^{\mathbb{Q}} \frac{1/D_1}{(1 + \beta)} \exp(r), \quad (2.3)$$

where r is the continuously compounded risk-free rate. Using our distributional assumptions, one can show that:

$$r = \ln \frac{1 + \beta}{\beta} + \mu - \sigma^2 \left(\frac{1}{2} + \frac{1}{\theta} \right).$$

We see that robustness does indeed affect the risk free rate: as θ decreases and the agent becomes more robust, there is downward pressure on r due to an increased precautionary savings motive. Notice that this effect is independent from the degree of the EIS: the increased precautionary savings motive allows us to derive a low risk free rate despite a unitary EIS. There is no upwards pressure on the risk free rate due to the usual substitution effect. Note also that when the agent does not seek robust policies ($\theta \rightarrow \infty$), we are back to the expected additive utility case.

In this context it is immediately possible to see how robustness considerations can alter the model's predictions regarding the equity premium. In this two-period example, the continuously compounded return on the risky asset is given by $\ln(D_1/S_0)$ and the observed equity premium is simply:

$$\mathbb{E} \ln \frac{D_1}{S_0} - r = \sigma^2 \left(1 + \frac{1}{\theta} \right).$$

Hence, it is theoretically possible to generate a high equity premium and a low risk free rate with low enough values of the robustness parameter θ . This result will be later confirmed and quantified in the context of our complete model.¹¹

¹⁰Miao (2004) also discusses a similar example.

¹¹Standard models with time additive expected utility violate the Hansen-Jagannathan (HJ) bound (Hansen and Jagannathan (1991)). Examining the HJ bound can shed more light on how robustness modifies a standard asset pricing framework. The robust Hansen-jagannathan bound needs to be modified and it takes the following

While the two-period example provides interesting insight into the implications of robustness, it does not allow us to analyze the main focus of this study: the term structure of interest rates. In the next section, we describe a continuous time, infinite horizon model in which we not only embed robustness considerations but also enrich the environment to allow for a time-varying investment opportunity set.

3 Robustness in a Continuous Time Model with Stochastic Volatility

In this section we present an infinite horizon, continuous time, general equilibrium model in which a robust representative agent derives optimal policies about consumption and investment. For simplicity we assume a Lucas tree type economy with a conditionally heteroskedastic growth rate of output. Our ultimate goal is to analyze the implied equilibrium yield curve in this economy and, in particular, identify the implications of robustness for the term structure of interest rates.¹²

form

$$\underbrace{\frac{\sigma(m)}{\mathbb{E}^{\mathbb{Q}}(m)} - \frac{h}{\sigma(R^e)}}_{\text{Model implied}} \geq \underbrace{\frac{|\mathbb{E}(R^e)|}{\sigma(R^e)}}_{\text{Data}}$$

where $m = \beta u'(C_1)/u'(C_0)$ is the intertemporal marginal rate of substitution (IMRS) and R^e is the excess return on a given asset class relative to the risk free rate (for a more general treatment see, for example, Barillas et al. (2007)). Why do we need to modify the HJ bound when the agent seeks robust policies? Note that the HJ bound links observed data (RHS) to the predictions of a candidate model (LHS). Since we refer to \mathbb{P} as the objective measure, the moments on the RHS are taken with respect to \mathbb{P} (i.e. the data are actually generated under \mathbb{P}). However, pricing is done using the subjective distribution \mathbb{Q} since the robust agent's IMRS serves as the pricing kernel. In our two period example, one can show that the LHS takes the following form

$$\frac{\sigma(m)}{\mathbb{E}^{\mathbb{Q}}(m)} - \frac{h}{\sigma(R^e)} \approx \sigma\left(1 + \frac{1}{\theta}\right)$$

Again, with the appropriate θ one can potentially satisfy the bound. Now the restriction is not only on the unconditional volatility of the IMRS as is usually the case with the HJ bound. The additional term $-h/\sigma(R^e)$ on the LHS can be very dominant, depending on the degree of robustness adopted by the agent.

¹²Gagliardini et al. (2004) also study robust control implications for the behavior of the term structure of interest rates in a Cox et al. (1985) type economy. We differ in two dimensions. They study a two factor model closely related to Longstaff and Schwartz (1992). We focus on a one factor model. More importantly, we study the empirical implications of our model and quantify the contribution of the state dependent market price of model uncertainty to our understanding of asset prices both in the equity and bond market. We also document supporting evidence to our key assumption of state dependent volatility in consumption growth. Finally, we estimate the implied degree of uncertainty aversion implied by the data.

3.1 The Economy

There is a single consumption good which also serves as the numeraire. We fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting a univariate Brownian motion $B = \{B_t : t \geq 0\}$. The diffusion of information is described by the filtration $\{\mathcal{F}_t\}$ on (Ω, \mathcal{F}) . All stochastic processes are assumed progressively measurable relative to the augmented filtration generated by B . Note however that this probability space corresponds to an approximating model and serves only as a reference point for the robust agent. The agent entertains a *set* of possible probability measures on (Ω, \mathcal{F}) , denoted by \mathcal{P} , which size is determined by a penalty function (relative entropy) imposed on the agent's utility function. Every element in \mathcal{P} is equivalent to \mathbb{P} (i.e., define the same null events as \mathbb{P}). We denote the distorted measure the agent chooses as $\mathbb{Q} \in \mathcal{P}$. The relative entropy imposes strong structure on the possible distorted measures. By Girsanov's theorem we require the distorted measure to be absolutely continuous with respect to the reference measure. Finally, the conditional expectation operator under \mathbb{P} and \mathbb{Q} is denoted respectively by $\mathbb{E}_t(\cdot) \equiv \mathbb{E}(\cdot | \mathcal{F}_t)$ and $\mathbb{E}_t^{\mathbb{Q}}(\cdot) \equiv \mathbb{E}^{\mathbb{Q}}(\cdot | \mathcal{F}_t)$.

Let D be an exogenous output process that follows a geometric Brownian motion and solves the following stochastic differential equation (SDE)

$$dD_t = D_t \mu dt + D_t \sqrt{v_t} dB_t \quad (3.1)$$

One can obviously think of D as a general dividend process of the economy. In either interpretation we will allow the trading of ownership on the tree that yields this output. The parameters μ and v are the local expectations (drift) and the local variance of the output growth rate, respectively. We assume that v follows a mean-reverting square-root process

$$\begin{aligned} dv_t &= (a_0 + a_1 v_t) dt + \sqrt{v_t} \sigma_v dB_t, \\ a_0 &> 0, \quad a_1 < 0, \quad \sigma_v \in \mathbb{R}, \quad 2a_0 \geq \sigma_v^2 \end{aligned} \quad (3.2)$$

Note that the same shock (Wiener increments) drives both the dividend growth rate and volatility processes.¹³ This assumption is made solely to retain the parsimonious description of the economy. The requirement $a_1 < 0$ guarantees that v indeed converges back to its steady state level $-\frac{a_0}{a_1}$ ($= \bar{v}$) at the velocity $-a_1$. This long run level is positive since $a_0 > 0$. The Feller condition $2a_0 \geq \sigma_v^2$ guarantees that the drift is sufficiently strong to ensure $v > 0$ a.e. once $v_0 > 0$. σ_v is a diffusion constant. Later we will show that the sign of σ_v plays an important

¹³We could also make the expected instantaneous output growth rate, μ , stochastic. By assuming, for example, an affine relation between μ_t and v_t , the model remains tractable and can be solved analytically. For the purpose of this paper, however, we maintain the assumption of a constant drift in the dividend process.

role in our model since it determines the risk exposure of default free bonds to the source of risk in the economy.

The assumption of stochastic volatility is important in our context and we will later document it empirically. We know that if v is constant, the market price of risk is also state independent, and thus the expectations hypothesis of the term structure of interest rates holds. This result stands in sharp contrast to the empirical evidence (e.g., Fama and Bliss (1987), Campbell and Shiller (1991), Backus et al. (1998), Cochrane and Piazzesi (2002)). We will show in the next section how stochastic volatility interacts with robustness considerations to affect the predictions of our model.

Let dR_t be the instantaneous return process on the ownership of the output process and S_t be the price of ownership at time t . Then, we can write

$$\begin{aligned} dR_t &\equiv \frac{dS_t + D_t dt}{S_t} \\ &= \mu_{R,t} dt + \sigma_{R,t} dB_t \end{aligned} \tag{3.3}$$

where μ_R and σ_R are determined in equilibrium. We also let r be the short rate process, which will be determined in equilibrium.

3.2 The Dynamic Program of the Robust Representative Agent

The robust representative agent consumes continuously and invests both in a risk-free and a risky asset. The risky asset corresponds to the ownership on the output process (the tree). The risk free asset is in zero net supply in equilibrium. In addition, as discussed earlier in Section 2, the agent chooses optimally a distortion to the underlying model in a way that makes his decisions robust to statistically small model misspecification. Formally, the agent has the following objective function

$$\sup_{C, \alpha} \inf_{\mathbb{Q}} \left\{ \mathbb{E}_t^{\mathbb{Q}} \left[\int_t^\infty e^{-\rho(s-t)} u(C_s) ds \right] + \theta \mathcal{R}_t(\mathbb{Q}) \right\} \tag{3.4}$$

subject to his dynamic budget constraint

$$dW_t = [r_t W_t + \alpha_t W_t (\mu_{R,t} - r_t) - C_t] dt + \alpha_t W_t \sigma_{R,t} dB_t \tag{3.5}$$

where \mathbb{Q} is the agent's subjective distribution, W is the agent's wealth, ρ is the subjective discount factor, C is the consumption flow process, α is the portfolio share invested in the risky asset, and θ is the multiplier on the relative entropy penalty \mathcal{R} , which will be interpreted

as the magnitude of the desire to be robust. When θ is set to infinity, (3.4) converges to the expected time additive utility case. A lower value of θ means the agent is more fearful of model misspecification and thus chooses \mathbb{Q} further away from \mathbb{P} in the relative entropy sense. In other words, the set \mathcal{P} is larger the smaller θ is.

Let \mathcal{L}_2 be the set of all progressively measurable univariate processes h such that $\int_0^\infty h_s^2 ds < \infty$ a.s.. Let \mathcal{H} be the set of all $h \in \mathcal{H} \subseteq \mathcal{L}_2$ such that the process $\xi^\mathbb{Q}$ defined by

$$\xi_t^\mathbb{Q} = \exp \left(\int_0^t h_s dB_s - \frac{1}{2} \int_0^t h_s^2 ds \right), \quad t \geq 0 \quad (3.6)$$

is a \mathbb{P} -martingale. Then, h defines the probability $\mathbb{Q} \in \mathcal{P}$ by $\mathbb{Q}(F) = \lim_{t \rightarrow \infty} E(1_F \xi_t^\mathbb{Q})$ for every $F \in \mathcal{F}$, and $\xi^\mathbb{Q}$ is also the conditional density process, or the Radon-Nikodym derivative of \mathbb{Q} with respect to \mathbb{P} , and satisfies

$$\xi_t^\mathbb{Q} = \mathbb{E}_t \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right), \quad t \geq 0$$

By Girsanov's theorem, for every $h \in \mathcal{H}$ we can define a Brownian motion under \mathbb{Q} as

$$B_t^\mathbb{Q} = B_t - \int_0^t h_s ds, \quad t \geq 0 \quad (3.7)$$

Using (3.7) we can also rewrite (3.6) as

$$\xi_t^\mathbb{Q} = \exp \left(\int_0^t h_s dB_s^\mathbb{Q} + \frac{1}{2} \int_0^t h_s^2 ds \right), \quad t \geq 0 \quad (3.8)$$

Note that $\xi^\mathbb{Q}$ is not a \mathbb{Q} -martingale.

With this setup at hand, the relative entropy process $\mathcal{R}(\mathbb{Q})$ for some $\mathbb{Q} \in \mathcal{P}$ can be expressed conveniently as¹⁴

$$\mathcal{R}_t(\mathbb{Q}) = \frac{1}{2} \mathbb{E}_t^\mathbb{Q} \left[\int_t^\infty e^{-\rho(s-t)} h_s^2 ds \right], \quad t \geq 0 \quad (3.9)$$

The expression in (3.9) allows us to rewrite (3.4) as

$$\sup_{C, \alpha} \inf_h \left\{ \mathbb{E}_t^\mathbb{Q} \int_t^\infty e^{-\rho(s-t)} \left[u(C_s) + \frac{\theta}{2} h_s^2 \right] ds \right\} \quad (3.10)$$

Note that now the infimization problem is well defined over \mathcal{H} .

Finally, using (3.7) we write (3.1), (3.2) and (3.5) under the distorted measure \mathbb{Q} . For

¹⁴See, for example, Hansen et al. (2006) and section 3, and especially proposition 4, in Skiadas (2003).

example, the wealth process under the agent's subjective distribution corresponds to

$$dW_t = \left[r_t W_t + \alpha_t W_t (\mu_{R,t} - r_t) - C_t + \underbrace{h_t \alpha_t W_t \sigma_{R,t}}_{\text{Drift contamination}} \right] dt + \alpha_t W_t \sigma_{R,t} dB_t^{\mathbb{Q}}. \quad (3.11)$$

In the context of the market return, for example, this drift contamination has an obvious interpretation: it is the uncertainty premium the agent requires for bearing the risk of potential model misspecification

$$dR_t = \left[\mu_{R,t} - \underbrace{(-h_t \sigma_{R,t})}_{\text{Uncertainty premium}} \right] dt + \sigma_{R,t} dB_t^{\mathbb{Q}}. \quad (3.12)$$

The process h is the (negative of) the process for the *market price of model uncertainty*. The diffusion part $\sigma_{R,t}$ on the return process is, as usual, the *risk exposure* of the asset. Their multiplication is the equilibrium *uncertainty premium*. In order to obtain the risk premium in the drift, one needs to rewrite the return process under the risk neutral measure. Let $\varphi \equiv \frac{\mu_R - r}{\sigma_R}$ be the local Sharpe ratio, or the process for the market price of risk in the model. Then, using the same arguments that lead to (3.7) we can link the reference measure to a risk neutral measure, denoted by B^q

$$B_t^q = B_t + \int_0^t \varphi_s ds, \quad t \geq 0 \quad (3.13)$$

or, alternatively, use both (3.7) and (3.13) to relate the risk neutral measure to the distorted measure

$$B_t^q = B_t^{\mathbb{Q}} + \int_0^t (\varphi_s + h_s) ds. \quad (3.14)$$

Then, the return process can be written as

$$dR_t = \left(\mu_{R,t} - \underbrace{\varphi_t \sigma_{R,t}}_{\text{Risk premium}} \right) dt + \sigma_{R,t} dB_t^q$$

Here, we see that the risk exposure σ_R is identical to the asset's uncertainty exposure. This leads to perfect correlation of risk and uncertainty premia in our model.

3.3 Optimal Policies with Robust Control

In this subsection we solve the robust representative agent's dynamic problem posited in Section 3.2. We use dynamic programming to derive closed form solutions for his optimal consumption and investment decisions policies together with the conditional distorted distribution.

Let $J(W_t, v_t)$ denote the agent's value function at time t where W_t and v_t correspond to current wealth and the conditional variance level respectively. The agent's Hamilton-Jacobi-Bellman (HJB) equation is^{15,16}

$$\rho J = \log C_t + \frac{\theta}{2} h_t^2 + \mathcal{D}^{\mathbb{Q}} J \quad (3.15)$$

where $\mathcal{D}^{\mathbb{Q}}$ is the Dynkin generator under the distorted measure. Informally, $\mathcal{D}^{\mathbb{Q}} J$ is $\mathbb{E}^{\mathbb{Q}}(dJ)/dt$ and is derived by applying Ito's lemma and using (3.11) and the distortion of (3.2) to characterize the dynamics of J . The only difference between this HJB equation and a standard one is the introduction of a *cost* and *benefit* for distorting the objective distribution. The cost is given by the relative entropy term $\frac{\theta}{2} h_t^2$ (pessimism is costly) and the benefit is hidden in the distortion of the Dynkin generator. The drift of the J process is distorted since the state processes are themselves distorted.

The solution for h from the infimization problem is given by

$$h_t = -\frac{1}{\theta} (J_{W,t} \sigma_{W,t} + J_{v,t} \sigma_v \sqrt{v_t}) \quad (3.16)$$

One can immediately see that the intuition from the 2-period example survives our infinite-horizon, continuous-time setting. First and foremost, the robustness correction h is state dependent. The robust agent derives the distorted conditional distribution in such a way that the reference conditional distribution first order stochastically dominates the chosen distorted conditional distribution. If it was not the case then there would be states of the world in which the robust agent would be considered optimistic. Also, the agent wants to maintain the optimal relative entropy penalty constant since θ is constant. In order to achieve this when conditional volatility is stochastic, the distortion has to be stochastic and increase with volatility (see expression (2.2)).

Second, the size of the distortion is inversely proportional to the penalty parameter θ : the distortion vanishes as $\theta \rightarrow \infty$. Third, whenever the marginal indirect utility and volatility of wealth (J_W and σ_W) are high, the agent becomes more sensitive to uncertainty and distorts

¹⁵See also Anderson et al. (2003) and Maenhout (2004) for similar formulations.

¹⁶See Appendix B for a more detailed derivation of the policies and the value function.

the objective distribution more. Low levels of wealth imply large marginal indirect utility of wealth. These are states in which the agent seeks robustness more strongly. The second term in the parentheses corresponds to the effect of the state v on the distortion h . Since $J_v < 0$ for all reasonable parametrizations, the sign of σ_v dictates the optimal response of the agent. Consider the benchmark case when σ_v is positive. Following a positive shock, marginal utility falls as consumption rises, and volatility v increases. Therefore, the investment opportunity set deteriorates exactly when the agent cares less about it. Since the evolution of v serves as a natural hedge for the agent, he reduces the distortion h . The opposite occurs when $\sigma_v < 0$.

Maximizing (3.15) over α , the optimal portfolio holding of the risky asset at time t can be expressed in two equivalent forms, each emphasizing a different aspect of the intuition. The first one is the myopic demand

$$\alpha_t = \frac{\mu_{R,t}^{\mathbb{Q}} - r_t}{\sigma_{R,t}^2}. \quad (3.17)$$

Equation (3.17) that the demand for the risky asset is myopic: the agent only cares about the current slope of the mean-variance frontier. However, this slope is constructed using his subjective beliefs. From an objective point of view, the agent deviates from the observed mean-variance frontier portfolio due to his (negative) distortion to the mean h : he believes the slope is lower and thus decreases his demand for the risky asset. The second form of the demand for the risky asset captures this idea

$$\alpha_t = \frac{\mu_{R,t} - r_t}{\sigma_{R,t}^2} + \frac{h_t}{\sigma_{R,t}}. \quad (3.18)$$

The first element on the right-hand side of equation (3.18) describes the myopic demand of a log-utility agent who is endowed with the objective measure. However, the pessimistic agent optimally reduces his holdings of the risky asset by $h/\sigma_R < 0$ since he believes the expected return on the risky asset is lower than the one implied by the objective measure.

We posit the guess that the value function is concave (log) in the agent's wealth and affine in the conditional variance

$$J(W_t, v_t) = \frac{\log W_t}{\rho} + \delta_0 + \delta_1 v_t \quad (3.19)$$

Now, we can use (3.19) to rewrite (3.16) as

$$h_t = -\frac{1}{\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right) \sqrt{v_t} \quad (3.20)$$

Here, we see that the distortion, or the (negative of the) market price of model uncertainty

is linear in the conditional volatility of the output growth rate. In equilibrium \sqrt{v} is the conditional volatility of the consumption growth rate.¹⁷

We can also rewrite (3.17) as

$$\alpha_t = \underbrace{\frac{1}{1 + \frac{1}{\rho\theta}} \frac{\mu_{R,t} - r_t}{\sigma_{R,t}^2}}_{\text{Myopic demand}} - \underbrace{\frac{\frac{1}{\theta}}{1 + \frac{1}{\rho\theta}} J_v \sigma_v}_{\text{Hedging demand}}$$

The first element on the RHS corresponds to a variant of the usual myopic demand for a risky asset in a log-utility setup. This term simply gives the trade-off between excess return compensation and units of conditional variance. Note that the coefficient is not unitary, as in the usual log problem. The reason is best understood when one keeps in mind the mapping between a robust control agent and an SDU agent with unitary EIS. When introducing robustness, we effectively increase risk aversion, but maintain the unitary EIS. This effect pushes down the demand schedule for the risky asset. The second element is the hedging-type component arising from uncertainty aversion, and it is larger in absolute terms the larger J_v or σ_v , ceteris paribus. The hedging part is positive since $J_v \sigma_v > 0$ due to the intuition given in (3.16).

The consumption policy is unchanged when the agent seeks robust policies: $C = \rho W$. The wealth and substitution effects still cancel out in our setup. When volatility increases, the agent decreases his holdings of the risky asset substantially since he cannot amortize the volatility increase through changes in his consumption. Unitary EIS implies a constant consumption-wealth ratio, thus all volatility changes are channelled through the asset market. In other words, robustness, or pessimism, entails that the agent perceives the local expectations on the risky asset to be lower than the objective drift on the same asset. The substitution effect implies that the agent should invest less since the asset is expected to yield low return in the future. In contrast, the wealth effect predicts that he should consume less today and save instead. In the

¹⁷It is possible to assume an exogenous process for μ separately from v and still maintain a fairly simple closed-form equilibrium. All one needs is to scale the local volatility of μ with the current level of \sqrt{v} . One such possible model will assume

$$d\mu_t = (x_0 + x_1 \mu_t) dt + \sqrt{v_t} \sigma_\mu dB_{2,t}, \quad x_1 < 0$$

Then, the value function for the robust agent is

$$J(W_t, \mu_t, v_t) = \frac{\log W_t}{\rho} + \delta_0 + \delta_1 \mu_t + \delta_2 v_t$$

and the robust correction is still linear in the conditional stochastic volatility

$$h = -\frac{\sqrt{v}}{\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_\mu + \delta_2 \sigma_v \right)$$

case of log utility, these two effects cancel each other. Consequently, the effect of robustness on the consumption policy is eliminated. Changing a log-agent's desire to be robust will only affects the risk free rate and the return on the risk free asset.

3.4 Robust Equilibrium

In this section we solve for the equilibrium price of the risky asset and the risk free rate. We define and discuss the implications of the robustness assumption on the equilibrium prices. Specifically, we will examine the level and volatility of both the equity premium and the risk free rate. First, we define a robust equilibrium:

Definition 1 *A robust equilibrium is a set of consumption and investment policies/processes (C, α) and a set of prices/processes (S, r) that support the continuous clearing of both the market for the consumption good and the equity market ($C = D, \alpha = 1$) and (3.10) is solved subject to (3.5), (3.2) and (3.7).¹⁸*

The only difference between this equilibrium definition and a conventional one is that the agent solves a robust control problem. We will now show that this affects the equilibrium short rate.

In equilibrium, since the agent consumes the output ($C = D$) the local consumption growth rate and the local output growth rate are the same ($\mu_C = \mu$). Also, the agent's equilibrium path of wealth is identical to the evolution of the price of the 'tree' since $\alpha = 1$. Therefore, $W = S$. Hence, $D = C = \rho W = \rho S$. As is usually the case with a log representative agent, not only the consumption wealth ratio is constant but so is the dividend-price ratio ($\frac{C}{W} = \frac{D}{S} = \rho$). We see that, as in the two-period example, the robustness considerations do not affect the consumption policy and the pricing of the 'tree'. In that case, what are the implications of the fact that the agent seeks robust policies? The effect shows up in the risk free rate and the way expectations are formed about growth rates or the return on the risky asset. The equilibrium risk free rate can be derived from (??)

$$\begin{aligned} r_t &= \rho + \mu_{C,t} + \sqrt{v_t} h_t - v_t \\ &= \rho + \mu_{C,t} - v_t \left[1 + \frac{1}{\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right) \right] \\ &= \rho + \mu - \phi v_t \end{aligned} \tag{3.21}$$

¹⁸The same definition also appears in Maenhout (2004). Without stochastic volatility considerations, he also derives the equilibrium risk free rate and equity premium.

For the remainder of the paper we define

$$\phi \equiv 1 + \frac{1}{\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right)$$

The usual comparative statics arguments apply to this short-rate equation. A higher subjective discount rate preference parameter ρ makes the agent want to save less, so that the equilibrium real rate must be higher to compensate the agent for saving as much as before. Higher future expected consumption growth makes the agent want to consume more today (substitution effect). The real rate must therefore be higher to prevent him from borrowing. Higher consumption volatility activates a precautionary savings motive, so that the real rate must be lower to prevent the agent from saving. The role of robustness can be interpreted in two ways. First, robustness distorts the expected consumption growth rate. Lower expected consumption growth rate lowers the equilibrium risk free rate since the substitution effect is now weaker. The second interpretation, which may be more intuitive, is that robustness amplifies the effect of the precautionary savings motive in the same direction ($h < 0$ when $\theta < \infty$), and thus lowers the equilibrium level of the short rate. All else equal, the robust agent wants to save more than an expected utility agent and therefore the former needs a stronger equilibrium disincentive to save in the form of lower risk free rate. In this latter interpretation, the distortion is proportional to consumption growth rate volatility and thus can be interpreted as a modification to the precautionary savings motive.

The equilibrium local expected return on the risky asset can immediately be derived from (3.3) and the fact that $S = D/\rho$

$$\begin{aligned} dR_t &= (\mu_{D,t} + \rho) dt + \sigma_{D,t} dB_t \\ &= (\mu_{D,t} + \rho + h_t \sigma_{D,t}) dt + \sigma_{D,t} dB_t^{\mathbb{Q}} \end{aligned}$$

And the observed equity premium is¹⁹

$$\begin{aligned} \mu_{R,t} - r_t &= \phi v_t = \underbrace{v_t}_{\text{Risk Premium}} + \underbrace{(\phi - 1) v_t}_{\text{Uncertainty Premium}} \\ &= \text{cov}_t \left(\frac{dC_t}{C_t}, dR_t \right) + \frac{1}{\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right) v_t \end{aligned}$$

¹⁹We use the qualifier ‘observed’ to emphasize again that what the agent treats as merely a reference model is actually the DGP. Therefore, anything under the reference measure is what the econometrician observe when he has long time series of data.

The equity premium has both a risk premium and an uncertainty premium components. The former is given by the usual relation between the agent's marginal utility and the return on the risky asset. If the correlation between the agent's marginal utility and the asset return is negative, the asset commands a positive risk premium $\left[cov_t \left(\frac{dC_t}{C_t}, dR_t \right) > 0 \right]$ and vice versa. The higher the degree of robustness (i.e., the smaller the parameter θ) higher are the uncertainty premium and the market price of model uncertainty. While a decrease in θ increases the equity premium, it also decreases the risk free rate through the precautionary savings motive. The EIS is independent of θ . By lowering θ we are effectively increasing the aversion to model uncertainty but not affecting the intertemporal substitution. Also, the distortion of equilibrium prices is not surprising since the agent believes consumption growth rate is lower than the actual growth rate under the reference model. Hence, his IMRS process is distorted.

We see that robustness can account for both a high observed equity premium and low level of the risk free rate. What about the volatility of the risk free rate? Since we do not change the substitution motive, the only magnification is through the precautionary savings. Empirically v is extremely smooth and, thus, contributes very little to the volatility of r .²⁰

Previous studies (e.g., Anderson et al. (2003), Skiadas (2003), Maenhout (2004)) have showed that without wealth effects, a robust control economy is observationally equivalent to a recursive utility economy in the discrete time case (Epstein and Zin (1989), Weil (1990)) or to a stochastic differential utility (SDU) in the continuous time economy as in Duffie and Epstein (1992a) and Duffie and Epstein (1992b). Thus, our combined market price of risk and uncertainty can be viewed as an effective market price of risk in the SDU economy.²¹ The difficulty with such approach is that it requires implausibly high degrees of risk aversion. Another difficulty arises in the context of the Ellsberg paradox. Our approach assumes that agents do not necessarily know the physical distribution and want to protect themselves against this uncertainty.

²⁰If we allow for a stochastic μ with positive correlation with v , fluctuations in v will be countered by movements in μ since they affect the risk free rate with opposite signs. In other words, if we allow the substitution effect and the precautionary motive to vary positively over time, the risk free rate can be very stable.

²¹Even though we do not lose the homotheticity of our problem since our agent has log preferences, Maenhout (2004) discusses the need to rescale the problem in order to obtain an exact mapping from the robust control economy to an SDU economy. We do not incorporate this rescaling since our interpretation focuses solely on an agent who faces Knightian uncertainty and acts as an ambiguity averse agent. Thus, we conduct this study with the intention of studying the behavior of both the market price of risk and uncertainty.

4 Pricing the Term Structure of Interest Rates

Denote the intertemporal marginal rate of substitution (IMRS) process by Λ where $\Lambda_t \equiv e^{-\rho t}/C_t$. Using Ito's lemma we characterize the dynamics of Λ

$$\frac{d\Lambda_t}{\Lambda_t} = -r_t dt - \sqrt{v_t} dB_t^{\mathbb{Q}} \quad (4.1)$$

where the drift is the (negative of) the short rate and the diffusion part is the market price of risk.

Using (4.1) it is straightforward to price default free bonds.^{22,23} We use the following guess for the functional form for the time t default-free zero-coupon bond price (an affine yield structure) that matures at time T . Let $\tau = T - t$

$$p(\tau; v_t) = \exp[\beta_0(\tau) + \beta_1(\tau) v_t]. \quad (4.2)$$

Start with the fundamental pricing equation where the expected marginal utility weighted price is a martingale

$$\mathbb{E}_t^{\mathbb{Q}}[d(\Lambda_t p_t)] = 0 \implies \mathbb{E}_t^{\mathbb{Q}}\left(\frac{dp_t}{p_t}\right) - r_t dt = -\frac{d\Lambda_t}{\Lambda_t} \frac{dp_t}{p_t}. \quad (4.3)$$

The excess expected return on a bond over the short rate is determined by the conditional covariance of the return on the bond and marginal utility, or alternatively, by the product of the market price of risk and the risk exposure of the bond. As usual, if they covary positively, the asset serves as a hedge against adverse fluctuations in marginal utility and commands a negative risk premium. In times of high volatility, the precautionary savings motive induces the agent to shift his portfolio away from the equity market and towards bonds. Such a shift induces an upward pressure on bond prices (and thus yields decrease). Therefore, bonds pay well in good times, rendering them a risky investment. Note, however, that the expectations are taken over the distorted measure. These distorted expectations affect prices in a systematic way relative to the prices that would have prevailed under the objective measure, introducing an uncertainty premium element into the price of the bond.

²²A more detailed derivation of the bond pricing rule, using the PDE approach, is in appendix C.

²³Our paper belongs to the vast literature on affine term structure models. The term structure literature is too large to summarize here but studies can be categorized into two strands - equilibrium and arbitrage free models. Our paper belongs to the former strand. The advantage of the equilibrium term structure models is mainly the ability to give meaningful macroeconomic labels to factors that affect asset prices. Dai and Singleton (2003) and Piazzesi (2003), for example, review in depth the term structure literature.

From (4.2) one can show that the risk premium on a default free bond is

$$-\frac{d\Lambda_t}{\Lambda_t} \frac{dp_t}{p_t} = \beta_1(\tau) \sigma_v v_t$$

where β_1 is positive and determines the cross section restrictions amongst different maturity bonds. The sign of the risk premium is determined by the correlation of the output growth rate and the conditional variance. In the next section we discuss the intuition behind the predictions of the model, and especially the role robustness plays in our context.

Moreover, the *observed* excess return that long term bonds earn over the short rate is not completely accounted for by the risk premium component. We derive the dynamics of the return on a bond with arbitrary maturity by applying Ito's lemma to (4.2). Under the objective measure we have,

$$\frac{dp(\tau; v_t)}{p(\tau; v_t)} = \left[r_t + \underbrace{\beta_1(\tau) \sigma_v v_t}_{\text{Risk Premium}} + \underbrace{\beta_1(\tau) \sigma_v v_t (\phi - 1)}_{\text{Uncertainty Premium}} \right] dt + \beta_1(\tau) \sigma_v \sqrt{v_t} dB_t.$$

In the presence of uncertainty aversion, there is an uncertainty premium that drives a wedge between the return on a τ -maturity bond and the short rate. The more robust the agent, the larger the market price of uncertainty is in absolute terms (i.e., ϕ is larger so $-h = (\phi - 1) \sqrt{v}$ is larger). Also, higher conditional variance increases the uncertainty premium since the agent distorts the mean of the objective model more. In other words, higher σ_v also increases the uncertainty exposure of the asset. We can express the uncertainty premium as

$$\underbrace{-h_t}_{\text{Price of uncertainty}} \times \underbrace{\text{diff} \left(\frac{dp(\tau; v_t)}{p(\tau; v_t)} \right)}_{\text{Uncertainty Exposure}} = \underbrace{\beta_1(\tau) \sigma_v v_t (\phi - 1)}_{\text{Uncertainty premium}}$$

where $\text{diff}(\cdot)$ is the diffusion part of the process. The intuition and implication of this result are discussed in the empirical section (5).

The yield on a given bond is simply an affine function of the conditional variance

$$\mathcal{Y}(\tau; v_t) = -\frac{1}{\tau} \ln p(\tau; v_t).$$

The two extreme ends of the yield curve are $\lim_{\tau \rightarrow 0} \mathcal{Y}(\tau; v_t) = r_t$ and $\lim_{\tau \rightarrow \infty} \mathcal{Y}(\tau; v_t) = \rho + \mu - a_0 \bar{\beta}_1$. Thus the spread is

$$\lim_{\tau \rightarrow \infty} \mathcal{Y}(\tau; v_t) - \lim_{\tau \rightarrow 0} \mathcal{Y}(\tau; v_t) = -a_0 \bar{\beta}_1 + \phi v_t$$

where the expression for $\bar{\beta}_1$ is given in Appendix C.

4.1 Why Can The Model Explain the Cross Section of Bond Yields? Some Intuition

In this section we explain more intuitively why the model accounts for the cross section regularities of bond yields. More importantly, we focus on the contribution of robustness considerations to the results.

4.1.1 Bond Returns and Upward Sloping Yield Curve

A bond price is the conditional expected IMRS. Ceteris paribus, a positive shock to the expected growth rate of consumption lowers the equilibrium bond price, and thus, increases the yield on that bond. The bond price decreases due to a negative substitution effect. If expected consumption growth rate has positive contemporaneous correlation with consumption, or negative marginal utility, the bond is considered a safe asset and therefore commands a negative risk premium. The opposite also holds true. Furthermore, the expected IMRS is also affected by the conditional variance of consumption growth but in the opposite direction. Holding everything else constant, a positive shock to the conditional variance of the growth rate of consumption increases the bond price, and thus, lowers the yield on that bond. Here, people want to save more due the precautionary savings motive and therefore, in equilibrium, bond prices are higher and yields are lower. Again, what determines the sign of the risk premium is the correlation of the conditional variance with marginal utility. If the correlation with marginal utility is negative the bond is considered a risky asset since it pays well in good times. Hence, investors require a positive risk premium on the bond.

Since the distortion h is linear in the conditional volatility of consumption growth (v), it is natural to think of robustness as magnifying the precautionary savings motive. Mean reversion of the conditional variance process coupled with a positive correlation between conditional variance and consumption growth entails a positive risk premium on long term bonds relative to short term bonds. Also, since long term yields are averages of future expected short term yields plus a risk premium, the average yield curve is expected to be upward sloping.

An alternative way of interpreting the average positive slope of the yield curve is by examining the objective and subjective (endogenous) evolution of the conditional variance of consumption growth rate. The (perceived) evolution of v under the distorted measure \mathbb{Q} is different from the evolution of v under the objective measure \mathbb{P} in two respects. Write (3.2)

under both measures

$$\begin{aligned} dv_t &= -\kappa_v (v_t - \bar{v}) dt + \sigma_v \sqrt{v_t} dB_t \\ &= -\kappa_v^{\mathbb{Q}} (v_t - \bar{v}^{\mathbb{Q}}) dt + \sigma_v \sqrt{v_t} dB_t^{\mathbb{Q}}. \end{aligned} \quad (4.4)$$

Here, κ_v is the velocity of reversion and \bar{v} is the steady state of v , both under the reference measure. However, the subjective velocity of reversion is

$$\kappa_v^{\mathbb{Q}} = \kappa_v - \sigma_v (1 - \phi) > \kappa_v \quad (4.5)$$

and the subjective steady state is

$$\bar{v}^{\mathbb{Q}} = \frac{\kappa_v}{\kappa_v^{\mathbb{Q}}} \bar{v} < \bar{v}. \quad (4.6)$$

Observation (4.6) is enough to explain the positive slope of the yield curve. Note that pricing is done using the IMRS of the robust agent and he thinks that the steady state of the conditional variance of consumption growth rate is lower than the objective target. By persistently missing the target, the agent on average believes that v is expected to decrease. In other words, he on average thinks that yields are expected to increase due to the effect of the precautionary savings motive on prices.

A different way of interpreting (4.4) is the following. The variance dynamics are characterized by a non-negative mean-reverting process. This process gravitates towards its steady state and the speed of reversion is stronger the further the variance level is from its steady state. Robustness introduces a negative distortion to the drift of the variance process ($h_t \sigma_v \sqrt{v_t} = \sigma_v (1 - \phi) v_t < 0$). A negative distortion to the drift that depends linearly on the level of the variance introduces zero as an additional focal point to the variance process. When the variance is above its objective steady state, both the distortion and the pull towards the objective steady state work in the same direction. However, when the variance is below its steady state, both forces work in opposite directions. The distortion always pulls down towards zero while the other force pulls the variance up towards its objective steady state. The point where these two forces are equal is the subjective steady state and it is between the objective steady state (positive) and zero, leading to (4.4).

4.1.2 Negative Contemporaneous Correlation Between the Spread and the Level of Yields (Yield Curve Rotation), and the Term Structure of Unconditional Volatilities of Yields

In quarterly data over the sample 52.Q2 – 06.Q4 the correlation between the level and slope of the real yield curve is -0.5083 with standard errors of 0.0992 (Newey-West corrected with 4 lags). Here, the slope is the difference between the 1-year and 3-months yields. This finding is robust over different time intervals and different frequencies. The model can account for this fact in the following way.²⁴ Recall that a positive shock to conditional volatility lowers yields. Also note that yields are perfectly (positively) correlated since all of them are an affine function of the same factor. However, short yields are more sensitive to conditional volatility shocks. To understand why, it helps to think about the mean reversion of the conditional variance (the ergodicity of its distribution). The effect of any shock is expected to be transitory. The full impact of the shock happens at impact and then the conditional variance starts reverting back to its steady state. Therefore, the effect of, say, a positive shock is expected to dissipate and yields are expected to start to climb back up. This expected effect is incorporated into long term yields immediately. Short yields in the far future are almost unaffected by the current shock since it is expected that the effect of the shock will disappear eventually. Since long term yields are an average of future expected short yields plus expected risk premia, they tend to be smoother than short term yields.

The expected risk premium is also a linear function of the state, and thus inherits its mean reversion. Therefore, the expected risk premium in the far future is also smoother than the risk premium in the short run. This also contributes to the rotation of the yield curve: since the short end of the yield curve is very volatile relative to the long end, whenever yields decrease, the spread increases (or become less negative, depending on the initial state). The opposite also holds true.

4.1.3 How Does the Model Account for the Rapidly Declining Slope of the Yield Curve and the High Persistence of Yields?

Traditionally, one factor models encounter an inherent difficulty in trying to account simultaneously for the rapidly declining slope of the yield curve (i.e., strong convexity of the slope of the yield curve) and the high persistence of yields. Time-series evidence implies that interest-rate shocks die out much more slowly than what is implied from the rapidly declining slope of the

²⁴We explain the intuition through the time variation of the conditional volatility of consumption growth rate. One can alternatively use the substitution channel and focus on time variation in expected consumption growth rate.

average yield curve (Gibbons and Ramaswamy (1993)).

Even though we present a one factor model, we can still account for these two facts with a single parametrization. The key lies in expression (4.5). The agent believes that the conditional variance reverts to its steady state *faster* than under the objective measure ($\kappa_v^Q > \kappa_v$). Since yields are affine functions of the conditional variance of consumption growth, they inherit the velocity of reversion of v under the objective model. In other words, the persistence of yields is measured ex-post and is solely determined by the objective evolution of v without any regard to what the agent actually believes.

At the same time, the slope of the yield curve (or the pricing of bonds) is completely determined by what the agent believes the evolution of v is. If κ_v^Q is substantially larger than κ_v , the slope of the yield curve can flatten at relatively short horizons, reflecting the beliefs of the agent that v will quickly revert to its steady state level. Since the agent persistently thinks that $\kappa_v^Q > \kappa_v$ the slope can be on average rapidly declining. When analyzing the results of our estimation we will show that this is indeed the case.

4.1.4 Biased Expectations: Pessimism and (the Reverse of) Doubt

Abel (2002) argues that one can potentially account for the equity premium and the risk free rate when modeling pessimism and doubt in an otherwise standard asset pricing (Lucas tree) model. Pessimism is defined as a leftward translation of the objective distribution in a way that the objective distribution first order stochastically dominates the subjective distribution. Doubt is modeled in a way that the subjective distribution is a mean preserving spread of the objective distribution.

There is evidence that people tend to consistently underestimate both market return and the conditional volatility of output growth rate (e.g., Soderlind (2006)). Also, Giordani and Soderlind (2006) confront the Abel (2002) suggestion with survey data and find strong support for the pessimism argument in growth rates of both GDP and consumption. The result is robust over forecasts of different horizon and with both the Livingston survey and the Survey of Professional Forecasters data. However, they also find evidence of overconfidence in the sense that forecasters underestimate uncertainty. Therefore, the evidence suggests the existence of the reverse of doubt.

Our model endogenously predicts both phenomena.²⁵ First, robustness requirements lead the agent to pessimistic assessments of future economic outcomes (e.g., expression (3.12) in

²⁵For a decision-theoretic link between ambiguity averse agent and the setup of Abel (2002), see Ludwig and Zimper (2006).

which the agent negatively distorts the expected return on the risky asset). Consequently, the agent persistently underestimates expected growth rates of both the risky asset and consumption. In that sense, robustness endogenizes the pessimism idea of Abel (2002). Our model also predicts biased expectations concerning the dynamics of the conditional variance process v in a way that is consistent with the data. Expressions (4.5) and (4.6) formalize this idea. In the case where $\sigma_v > 0$ (an assumption that we later support empirically), a pessimistic assessment of expected output growth rate leads to what can be interpreted as optimistic beliefs about future output growth volatility. In other words, the model predicts also the reverse of doubt. Note that here the agent knows exactly the current conditional variance but wrongly estimates its future evolution.

5 The Empirical Study

In this section we undertake three tasks. First, we provide empirical support for our assumption that the volatility of consumption growth is state dependent. Our discussion complements the analysis of Bansal and Yaron (2004) and Bansal et al. (2005) who argue that there is stochastic volatility in the growth rate of consumption. Second, we estimate our model.²⁶ There are six parameters in the model, five of which are standard. Third, we interpret the non-standard parameter θ using *detection error probabilities* to map θ . Since the model is a description of a real economy, all the data we use are expressed in real terms. The description and discussion of the data are relegated to Appendix D.²⁷

5.1 Conditionally Heteroskedastic Consumption Growth

In this subsection we provide direct empirical evidence about the level and behavior of the conditional variance of real aggregate consumption growth. We examine two measures of conditional volatility: realized volatility and series estimated from various GARCH specifications.

5.1.1 ARMAX-GARCH Real Consumption Growth Rate

We start with a simple univariate time series parametric estimation. The model we are fitting to the consumption growth process is an ARMAX(2, 2, 1) model and a GARCH(1, 1) to the

²⁶Wachter (2001), for example, studies the effect of consumption externalities (habits) on the term structure of interest rate by drawing empirical restrictions from consumption data and both the equity and bond markets.

²⁷A few studies, for example Brown and Schaefer (1994) and Gibbons and Ramaswamy (1993), also use real data to estimate a term structure model. However, they do not draw restrictions from the equity market and consumption data and their preferences assumption is standard which implies that the equity premium and risk free rate puzzles are still present in the models they estimate.

innovations process:

$$\begin{aligned}
A(L) \frac{\Delta C_t}{C_{t-1}} &= c + B(L) R_{t-1} + C(L) \eta_{C,t}, \\
\eta_{C,t+1} &= \sigma_{C,t} \varepsilon_{C,t+1}, \quad \varepsilon_{C,t} \sim N(0, 1), \\
D(L) \sigma_{C,t} &= \omega + F(L) \eta_{C,t}^2,
\end{aligned} \tag{5.1}$$

where A, B, C, D, F are polynomials of orders 2, 1, 2, 1, 1 respectively, in lag operators. $\frac{\Delta C_t}{C_{t-1}}, R_t, \eta_t$ are, respectively, the realized real consumption growth rate at time t , the real return on the aggregate market index at time $t - 1$, and an innovation process with time-varying variance. In Figure 5.1 we plot the GARCH volatility estimates for both real aggregate consumption growth rate and the real return on the aggregate stock market. We also plot a measure of realized volatility for both consumption growth and market return series that we obtain by fitting an ARMA(2, 2) to the original data and then use the square innovations to construct the realized variance series. The sample period is $Q2.52 - Q4.06$.

First, there seems to be evidence of what has been dubbed as the ‘Great Moderation’ (e.g., Stock and Watson (2003)). It is clear that consumption growth volatility has slowly declined over the sample period but the volatility of the market return did not. This pattern is apparent in both measures of conditional volatility. Second, it seems that there are both high frequency (business cycle) fluctuations and a very low frequency stochastic trend in consumption growth volatility. We will show that the estimation procedure mostly identifies these higher frequency movements in the conditional variance and not the very low frequency movements. Our hypothesis is that higher frequency fluctuations are channeled through the asset market while there are other aspects which we do not identify that contribute to the low frequency fluctuations. In other words, when we estimate the full model, the effect of the equity and bond market restrictions is reflected in the implied persistency of the conditional variance process. Here, we use the Hodrick-Prescott filter with parameter 1600 to disentangle these two components of consumption growth volatility. Figure 5.2 presents this result and makes clear that the decline in the low frequency component started in the ‘60, before the Great Moderation.²⁸

²⁸In our model it is hard to make ‘conditional’ statements about the economy, mainly because we modeled a constant drift to the consumption growth rate process. It is obviously interesting to think about the correlation structure of expected consumption growth rate and the conditional variance process. Empirically, there is evidence that suggests that interest rates are procyclical (e.g., Donaldson et al. (1990)) and volatility is either countercyclical or at least slightly leads expected growth rates which are believed to be countercyclical (e.g., Whitelaw (1994)). Our conditional variance process is assumed to correlate positively with realized consumption growth rate. Also, the conditional variance correlate negatively with interest rates. In this sense, variance and real interest rates behave as in the data. If, for example, expected growth rate correlate negatively with realized consumption growth rates, they will correlate negatively with the conditional variance. In that case, a positive

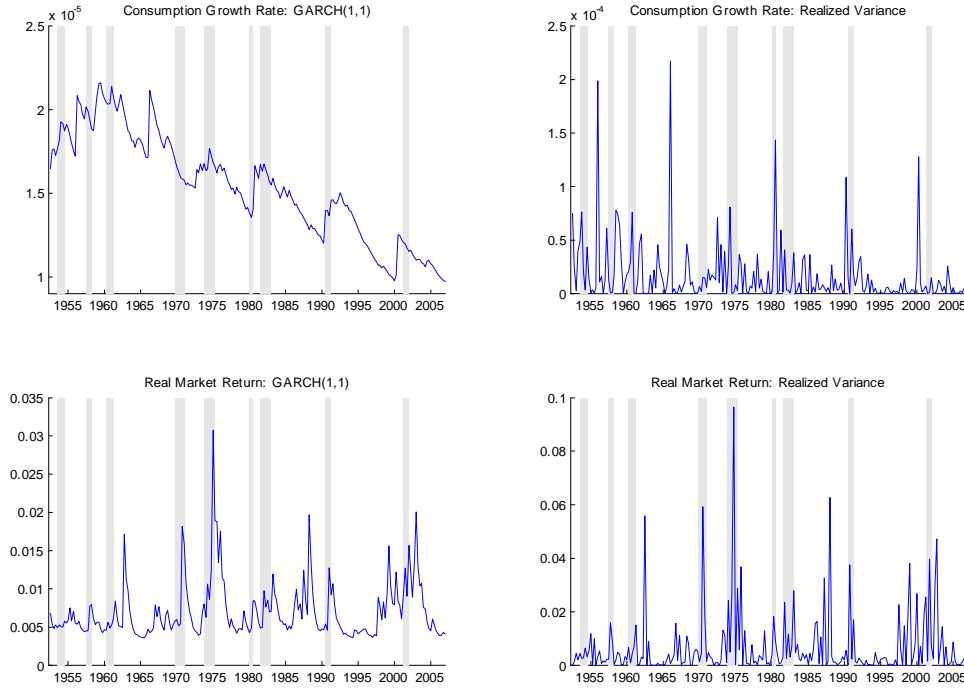


Figure 5.1: ARMAX-GARCH estimation for both real consumption growth rate and real aggregate market return. We fit model (5.1) and present the GARCH estimates for the conditional variance of real consumption growth rate and real aggregate market return in the left panel. The right panel present the square innovations from an ARMX specification to real consumption growth and real aggregate market return. The quarterly data is $Q2.52 - Q4.06$. The gray bars are contraction periods determined by the NBER.

We also use the volatility estimates to explain asset prices (see also Chapman (1997), Bansal and Yaron (2004), Bansal et al. (2005)). In particular, in figure 5.3 we examine the dynamic cross correlation patterns between consumption growth volatility obtained from the GARCH estimation in (5.1) and the spread between the real 1-year real yield and the real 3-months real yield.

These patterns agree with the model's predictions. We know that shorter maturity yields respond more than longer maturity yields to a volatility shock. This result is mainly due to the ergodicity of the state variable that affect yields. If the state is assumed to revert back to a known steady state, we expect the longer yield to have a smaller response to contemporaneous

shock to consumption growth rate will have a double negative effects on real interest rates. Expected growth rates will be low and thus the substitution effect will make equilibrium real interest rates lower. At the same time, conditional variance will be higher and the precautionary savings motive will push the equilibrium real interest rate even lower. Also, Chapman (1997) documented the strong positive correlation of real yields and consumption growth rate when excluding the monetary experiment period of 1979 – 1985.

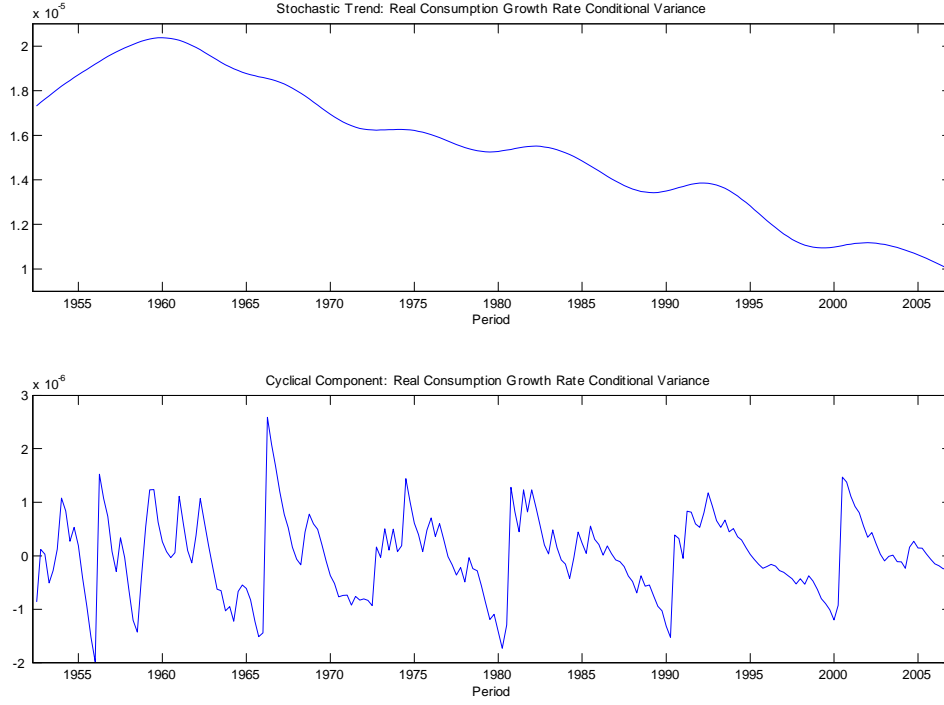


Figure 5.2: HP-filtered conditional variance of real consumption growth rate derived from an ARMAX-GARCH estimation in (5.1). The top panel presents the low frequency trend and bottom panel presents the cyclical component. The HP-filter parameter is 1600. The quarterly data is over the period $Q2.52 - Q4 - 06$.

shocks. Note that we do not identify the type of shock in this exercise. We merely observe a shock that happens to affect both consumption growth volatility and the bond market.

The second result is the sign response of the yields to a volatility shock. When conditional volatility increases we see that yields decrease. From the precautionary savings motive effect we do expect such response. Since in our model ambiguity aversion amplifies the precautionary savings motive, we expect this channel to play an important role when linking consumption growth volatility and yields. When combining these two results, we expect the spread to increase with a volatility shock. In other words, on average, the yield curve rotates when a shock to volatility occurs.

There are three caveats to these results. First, the upper left panel in Figure 5.1 depicts the behavior of the conditional variance of real consumption growth. One can argue that the series exhibit a non-stationary behavior. If this is the case, then the GARCH process is potentially misspecified. Given the slow-moving component we identified, it is hard to convincingly argue against such hypothesis. Second, our macro data is sampled at quarterly frequency. Drost and

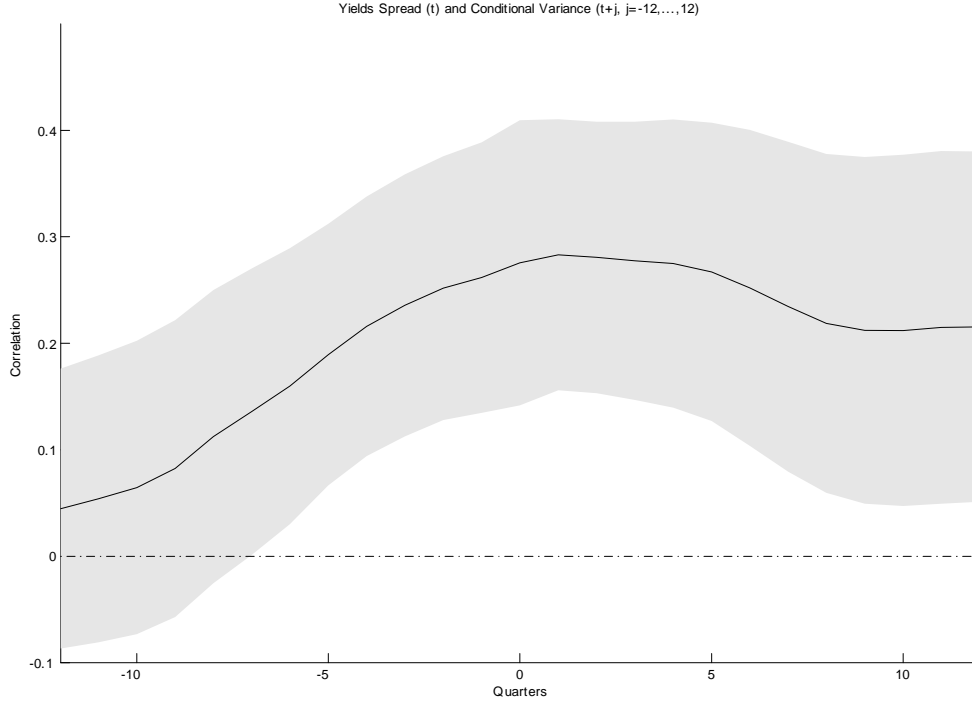


Figure 5.3: Dynamic cross-correlation between real consumption growth rate volatility and the real spread between the 1 year and 3 months yields. The quarterly data covers the period $Q2.52 - Q4.06$.

Nijman (1993) have shown that temporal aggregation impedes our ability to detect GARCH effects in the data. Even if our model is not misspecified, the fairly low frequency sampling may suggest it is (see also Bansal and Yaron (2004)). Third, we showed that the (sign of the) correlation between shocks to realized consumption growth and the conditional variance is important in explaining risk and uncertainty premia. The simple GARCH exercise does not help us identify the sign of this correlation. We address this difficulty next.

5.1.2 Real Dividends Growth Rate: GJR-GARCH

Since we argue that the sign of σ_v plays an important role in understanding risk premia in our model, we also estimate a GJR-GARCH(1,1) (Glosten et al. (1993)). Originally, this model was constructed to capture ‘leverage’ effects when examining market returns (i.e., a negative shock to returns means lower prices and more leveraged firms, hence higher volatility of future returns). Here we use it with a different interpretation in mind. We use the leverage coefficient to extract information about the sign of the correlation between consumption/dividends growth rate innovations and conditional variance innovations. Since we argue that the sign of σ_v is

positive, as indicated by asset prices behavior, we hope to find the reverse of a leverage effect.²⁹ We fit the following time series model

$$\begin{aligned}\frac{\Delta C_t}{C_{t-1}} &= c + \eta_{C,t}, \\ \eta_{C,t+1} &= \sigma_{C,t} \varepsilon_{C,t+1}, \quad \varepsilon_{C,t} \sim N(0, 1), \\ D(L) \sigma_{C,t} &= \omega + F(L) \eta_{C,t}^2 + G(L) I_{\{\eta_{C,t} < 0\}} \eta_{C,t}^2,\end{aligned}\tag{5.2}$$

where the polynomial G captures the leverage effects and

$$I_{\{\eta_{C,t} < 0\}} = \begin{cases} 1 & \eta_{C,t} < 0 \\ 0 & \text{otherwise} \end{cases}.$$

We regress the realized consumption growth rate only on a constant (effectively demeaning the growth rate) since we assume in our model that dividends growth rate drifts on a constant. The more negative η is, the larger is η^2 . Thus, we expect the leverage effect coefficient to be negative in order to capture the positive correlation between shocks to growth rates and conditional variance. In most lag specifications we estimated, the leverage coefficients in the G polynomial have a negative sign, which suggests that negative shocks to the dividends growth rate implies a negative shock to the conditional variance. However, and perhaps not surprisingly, with quarterly frequency data it is hard to detect these GARCH effects. Leverage effects are especially hard to detect. In most cases we cannot reject the null that leverage effects are not present. In order to investigate the sign of σ_v further, we use real dividends instead of consumption. To alleviate the problem with the GARCH estimation, we use *monthly* data.³⁰ Figure 5.4 displays the results of a GJR-GARCH(1, 1) estimation where c is the unconditional mean of the real growth rate of aggregate dividends

This figure shows the presence of volatility clustering. The estimation procedure suggests that σ_v is indeed positive since the leverage coefficient is *always* negative and statistically significant. On average, when a negative shock hits the dividends growth rate, we tend to see a decline in the conditional variance of the same process. Table 5.1.2 summarizes the estimation results for the leverage coefficient over different time intervals.³¹

It is interesting to note that the earlier post-war data supports more strongly the hypothesis that shocks to dividends are positively correlated with shocks to volatility. This covariation

²⁹Even though our interpretation has nothing to do with the leverage effect discussed in Glosten et al. (1993), we still use this term for convenience.

³⁰We obtained the real dividends series from Robert Shiller's website. See also Appendix D.

³¹This suggestive evidence is also consistent with different time intervals and with EGARCH estimation (see Nelson (1991)) over the same time intervals. Results are available from the authors upon request.

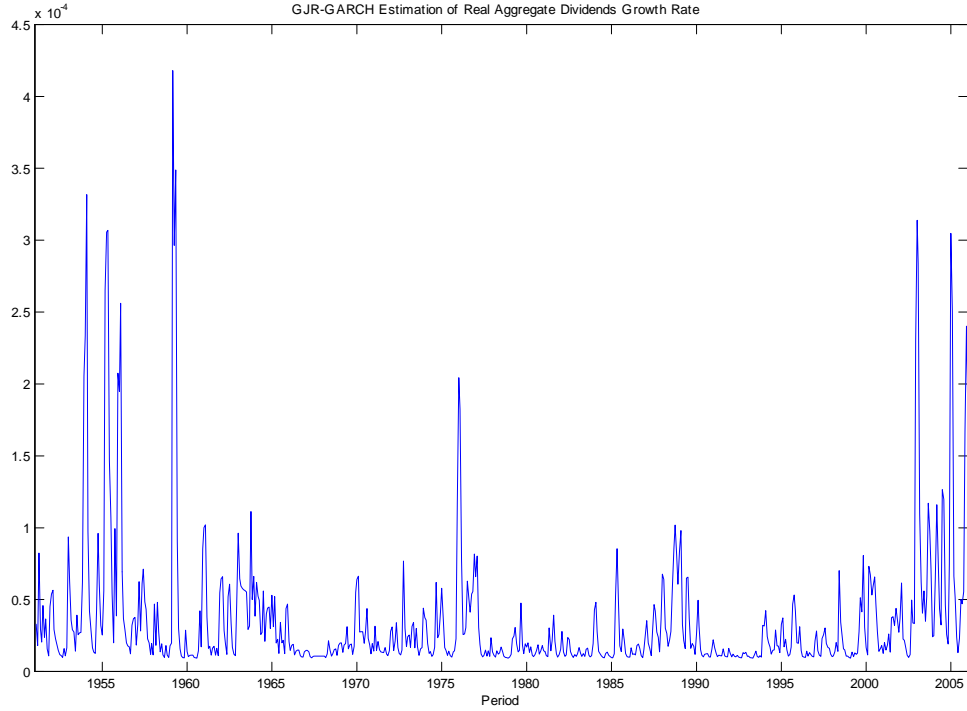


Figure 5.4: GJR-GARCH(1, 1) estimation (model 5.2) of the conditional variance of real aggregate dividends growth rate with monthly observations over the period $M1.52 - M4.06$.

measures the risk exposure of default free bonds to risk and uncertainty. If the market prices of these risks and uncertainty did not move in the opposite direction one should, *ceteris paribus*, expect to observe higher risk premia in the earlier part of the sample.

In summary, the data seems to confirm two things. First, the existence of a small time-varying component in the volatility of growth rates. Second, the correlation of shocks to dividends growth rate and shocks to conditional variance is positive.

5.2 Model Estimation

In this section we present and interpret our complete model estimation results. Since the model permits closed-form expressions for first and second moments we use the generalized method of moments (GMM) in the estimation procedure (Hansen (1982)). Even though conditional variance is not directly observable in the data it is theoretically an affine function of the short rate (or any other real yield with arbitrary maturity). Therefore, we use the short rate as an

Table 5.1: Estimating the ‘leverage’ coefficient over different time intervals. The data is monthly real aggregate dividends over $M1.52 - M12.06$ from Robert Shiller’s website. A negative point estimate means that a negative shock to realized dividends growth rate is accompanied by a negative shock to the conditional variance of dividends growth rate.

Period	‘Leverage’ Coefficient	Standard Errors
$M1.52 - M12.06$	-0.390	0.129
$M1.62 - M12.06$	-0.242	0.175
$M1.72 - M12.06$	-0.312	0.215
$M1.82 - M12.06$	-0.263	0.163
$M1.90 - M12.06$	-0.256	0.195
$M1.52 - M12.81$	-0.509	0.167
$M1.52 - M12.89$	-0.442	0.156

observable that completely characterizes the behavior of the conditional variance.³² We also compare the moments implied by the model to their empirical counterparts.

5.2.1 Orthogonality Restrictions and Identification

Our procedure is similar to the one used by, for example, Chan et al. (1992). Our approach is to focus mainly on the time series restrictions to estimate the structural parameters. We do not focus on the cross sectional restrictions of the model as in Longstaff and Schwartz (1992) and Gibbons and Ramaswamy (1993). Since we have a single factor model, yields are perfectly correlated. Therefore, including cross sectional restrictions may reduce the power of the overidentifying restrictions in small samples. We use our point estimates to generate the model’s implied yield curve and compare it to the empirical yield curve. In that sense, our approach is more ambitious. It is important to note that since our model only makes statements about the real economy, all the data we use is denominated in real terms. Other authors have used nominal data to estimate real models (e.g., Brown and Dybvig (1986)).

We need to estimate 6 parameters $\{a_0, a_1, \mu, \rho, \theta, \sigma_v\}$. We form orthogonality conditions

³²We also used the simulated method of moments (SMM, Duffie and Singleton (1993)) to estimate the model. This method is natural when the model contains unobservables. The results we obtain using SMM support the results we obtain using GMM and are available from the authors upon request. Bansal et al. (2007) also compare their GMM estimates to an SMM estimates and conclude that in the presence of time averaging, using SMM can prove useful.

implied by the model using the following notation

$$\begin{aligned} Y_{t+1} &\equiv \left[\Delta \mathcal{Y}(1; v_{t+1}), R_{t+1}, \frac{\Delta C_{t+1}}{C_t}, \mathcal{Y}(4; v_t) - \mathcal{Y}(1; v_t) \right], \\ X_t &\equiv \mathcal{Y}(1; v_t), \\ Z_t &\equiv \left[1, \mathcal{Y}(1; v_t), R_t, \frac{\Delta C_t}{C_{t-1}} \right], \end{aligned}$$

where Y_{t+1} is observed at time $t + 1$ and contains the change in the one-quarter real yield ($\Delta \mathcal{Y}(1; v_{t+1})$), the realized real aggregate market return (R_{t+1}), the realized real aggregate consumption growth rate $\left(\frac{\Delta C_{t+1}}{C_t}\right)$ and the real spread between the 1-year and 3-months real yields ($\mathcal{Y}(4; v_t) - \mathcal{Y}(1; v_t)$). X_t is the explanatory factor. We use the 3-months yield as a sufficient statistic for the unobserved conditional variance process. Last, we use lagged 3-months, market return and realized consumption growth rate as instruments in the vector Z_t .

The stacked orthogonality conditions are given in m

$$\begin{aligned} u_{1,t+1} &\equiv Y_{t+1} - \mu_{Y,t}|X_t, \\ u_{2,t+1} &\equiv \text{diag} \left(u_{1,t+1} u'_{1,t+1} - \sigma_{Y,t} \sigma'_{Y,t} | X_t \right), \\ m_{t+1} &\equiv \begin{bmatrix} u_{1,t+1} & u_{2,t+1} \end{bmatrix} \otimes Z_t. \end{aligned}$$

We draw first and second moment restrictions. $\mu_{Y,t}|X_t$ and $\sigma_{Y,t}|X_t$ have the parametric forms implied by the model and are affine in X_t .

What about identification? Note that since robust preferences are observationally equivalent to recursive preferences, disentangling the risk aversion coefficient from the robustness parameter θ is generally not trivial. Since we have log-utility we do need to worry about such a potential identification problem: log preferences restrict to unity the EIS and risk aversion and thus allow us to identify the uncertainty parameter θ . Also, μ is identified through the consumption growth rate restriction. Once μ is identified, we can identify ρ from the aggregate market return condition. The three parameters that govern the dynamics of the conditional variance v can be identified either from the second moment of consumption growth rate or the second moment of the aggregate market return. Also, the bond market contributes important information about v . The fact that identifying the dynamics of v is done through these three channels can potentially create some ambiguity in the interpretation of the level and speed of reversion of the conditional variance. Nevertheless, we believe that these sources of information shed some new light on the dynamics of v in a way that will be clear in our interpretation of

the point estimates, a task we undertake next.

5.2.2 Point Estimates of Structural Parameters

Table 5.2 presents the point estimates over different time periods. In Table 5.3 we perform the same estimation exercise without including the volatility of consumption growth rate in our set of moments.

Table 5.2: Model estimation with consumption volatility over different time intervals. The data is in quarterly frequency and in quarterly values. 6 parameters are estimated using an iterated GMM. There are 8 moments and 4 instruments that produce 32 orthogonality conditions. T is the number of observation in each estimation. Robust t-statistics are indicated below each point estimate. The standard error are corrected using the Newey-West procedure with 4 lags. p-val is the p-value for the J-test statistic distributed χ^2 with 26 degrees of freedom. The DEP column reports the detection error probabilities.

Period	T	a_0	a_1	μ	ρ	θ	σ_v	J-stat/P-val	DEP
$Q2.52 - Q4.06$	218	0.0005	-0.1951	0.0052	0.0145	9.5730	0.0189	33.5147	4.41%
		6.7904	-10.1242	29.0936	7.6824	3.7579	7.7160	14.77%	
$Q1.62 - Q4.06$	180	0.0004	-0.1594	0.0051	0.0125	14.9578	0.0175	28.0792	11.01%
		6.9588	-10.7284	28.2098	6.3202	2.7028	5.5972	35.46%	
$Q1.72 - Q4.06$	140	0.0004	-0.1611	0.0048	0.0148	11.8660	0.0149	22.3470	10.86%
		6.3636	-10.3456	28.4321	6.1634	2.5565	5.1858	66.96%	
$Q1.82 - Q4.06$	100	0.0004	-0.1314	0.0054	0.0208	6.6591	0.0082	17.3000	7.31%
		6.2872	-6.6014	34.0290	8.7880	3.7862	7.4882	89.97%	
$Q1.90 - Q4.06$	68	0.0003	-0.0997	0.0050	0.0172	9.0779	0.0064	12.2261	14.49%
		8.1423	-9.7387	36.4052	10.5310	5.7235	10.3818	98.98%	
$Q2.52 - Q4.81$	118	0.0007	-0.2528	0.0050	0.0115	13.4790	0.0431	20.2067	17.11%
		7.1360	-13.8242	20.8878	5.6148	2.7839	6.0434	78.17%	
$Q2.52 - Q4.89$	149	0.0006	-0.2232	0.0053	0.0140	10.6126	0.0290	24.2714	10.10%
		6.2481	-9.5348	24.3649	6.3682	3.2241	6.9336	56.04%	

Table 5.3: Model estimation without consumption volatility over different time intervals. The data is in quarterly frequency and in quarterly values. 6 parameters are estimated using an iterated GMM. There are 8 moments and 4 instruments that produce 32 orthogonality conditions. T is the number of observation in each estimation. Robust t-statistics are indicated below each point estimate. The standard error are corrected using the Newey-West procedure with 4 lags. p-val is the p-value for the J-test statistic distributed χ^2 with 26 degrees of freedom. The DEP column reports the detection error probabilities.

Period	T	a_0	a_1	μ	ρ	θ	σ_v	J-stat/P-val	DEP
$Q2.52 - Q4.06$	218	0.0011	-0.1903	0.0052	0.0159	23.5142	0.0331	26.4281	15.56%
		6.0765	-9.0950	25.1487	7.0689	3.1259	6.7880	23.38%	
$Q1.62 - Q4.06$	180	0.0009	-0.1666	0.0051	0.0134	39.9255	0.0335	25.7987	24.97%
		5.6809	-10.0265	25.9144	5.9769	2.1620	5.3863	26.04%	
$Q1.72 - Q4.06$	140	0.0010	-0.1719	0.0048	0.0157	31.0175	0.0286	20.3411	24.16%
		4.8876	-9.1749	24.2349	5.9522	2.0961	5.0048	56.17%	
$Q1.82 - Q4.06$	100	0.0008	-0.1336	0.0054	0.0214	15.9299	0.0159	16.4269	18.93%
		5.2730	-6.4625	31.9802	7.9282	2.9953	6.8526	79.42%	
$Q1.90 - Q4.06$	68	0.0006	-0.1006	0.0050	0.0172	22.3478	0.0125	11.7640	27.01%
		7.4532	-7.9747	32.3043	9.4070	4.0868	9.7618	96.21%	
$Q2.52 - Q4.81$	117	0.0013	-0.2332	0.0050	0.0119	34.7236	0.0749	18.0621	30.07%
		6.6457	-10.2125	18.5634	5.4752	2.3172	5.5378	70.23%	
$Q2.52 - Q4.89$	149	0.0012	-0.2069	0.0053	0.0148	26.6319	0.0530	20.8378	23.44%
		5.5588	-8.0626	21.2137	5.9463	2.7119	6.3294	53.08%	

Aside from the robustness parameter θ , all coefficients are immediately interpretable. All parameters are statistically different from zero. Also, the model is not being rejected according to the J-test. We will explain the DEP's column later.

Note that μ is stable and equal to the average real aggregate quarterly consumption growth rate over the sample. Similarly, ρ is stable over different samples and invariant to the consumption volatility restriction.

One obvious finding is that the estimated θ is sensitive to the inclusion of consumption volatility in the estimation. When the volatility of consumption growth is ignored, the procedure is not restricted by the smooth consumption process and thus the implied pricing kernel (SDF) is much more volatile and more robustness is not needed to justify the observed asset prices. In this sense, the implied volatility of the SDF is closer to the Hansen-Jagannathan bound. Also, note that a_0 and σ_v are much larger when we do not impose the consumption volatility restriction. The reason is that the procedure mainly picks up the aggregate market return volatility, which is much larger than the volatility of consumption growth rate. The implied evolution of v is much more volatile when consumption growth volatility is excluded.

Interestingly, the velocity of reversion ($-a_1$) of v is invariant to the consumption growth rate volatility. What is obvious from Tables 5.2 and 5.3 is that the estimation procedure detects mostly high frequency movements and not the potential slow moving component in consumption growth volatility we identified earlier (Figure 5.2). Hence, it appears that the high-frequency component from the market data dominates in the full-model estimation.

Panel A of Table 5.2.2 presents the half life of the volatility shock process implied by the estimation procedure. We also present in that panel the perceived half life by the robust agent. Expression (4.5) shows that the perceived velocity of mean reversion is faster than the physical speed at which shocks to volatility die out. In general, the point estimates imply that shocks to volatility die out relatively fast. For comparison purposes, Panel B of Table 5.2.2 presents the implied reversion coefficient and half life derived using the autoregressive coefficient we calculated from the GARCH estimated conditional variance series in (5.1) without adding the market as an explanatory variable to the consumption growth rate.³³ These results confirm that without forcing asset market restrictions on the consumption series, we observe a very slow moving process for conditional variance. At the same time, the conditional variance of the aggregate market return is much less persistent. The general estimation procedure results in

³³Our point estimates correspond to quarterly data. In general, with data sampled at quarterly frequency one can map an autoregressive coefficient to a coefficient governing the speed of reversion as our κ_v . Let $\hat{\alpha}$ denote the autoregressive coefficient. Then, the quarterly speed of reversion coefficient $\kappa_v = -\ln(\hat{\alpha})$ and the half life is $\ln(2)/\kappa_v$.

panel A are, to some extent, a combination of these two effects.³⁴

Table 5.4: Panel A: Point estimates of the velocity of reversion coefficient and the implied half life (in quarters) of the conditional variance process. Objective refers to the physical rate in which the conditional variance gravitates to its steady state. Distorted refers to the rate in which the robust agent believes the conditional variance gravitates to its steady state. These point estimates are from the estimation procedure that imposes the volatility of real aggregate consumption growth rate as a moment condition. Panel B: implied reversion coefficients and half lives (in quarters) for the conditional volatility of consumption growth rate and aggregate market return derived from the GARCH procedure. The consumption growth rate mean is modeled as an ARMAX(2,2,1) and the aggregate market return is modeled as ARMA(2,2).

Panel A:	$Q2.52 - Q4.06$		$Q1.90 - Q4.06$	
	Estimate	Half Life (Q)	Estimate	Half Life (Q)
Objective	0.1951	3.553	0.0997	6.952
Distorted	0.2994	2.315	0.1343	5.161
Panel B:	$Q2.52 - Q4.06$			
	Estimate		Half Life (Q)	
Consumption	0.010		68.373	
Market	0.4069		1.704	

Interestingly, in our benchmark estimation result we find the half life of the conditional variance process to be 3.553 quarters. The recent ‘long run risks’ literature usually calibrates asset pricing models with a highly persistent conditional variance process.³⁵ For example, Bansal and Yaron (2004) assume that the autoregressive coefficient (with monthly frequency data) in the conditional variance of the consumption growth process is 0.987.³⁶ This number implies a half life of 13.24 quarters, which is almost 4 times higher than the number we obtain in our empirical results. As explained earlier, this difference is driven largely by the inclusion of equity and bond markets in our set of moments. What we show in this paper is that robust decision making coupled with state dependent volatility requires moderate levels of persistence in the conditional variance of the consumption growth process. Recall that we assume a constant drift in consumption growth. If we assume a stochastic and highly persistent μ , as in Bansal and Yaron (2004), we would need to worry about the volatility of the risk free rate. In other words, if the substitution effect channel is very persistent and the precautionary savings motive

³⁴We conduct this comparison only for the entire period $Q2.52 - Q4.06$ since we want to examine evidence concerning very low frequency components. Even our longest sample is somewhat short to conveniently detect the slow moving component. We believe that shorter samples will make the detection exercise impossible.

³⁵Bansal and Yaron (2004) find that introducing a small highly persistent predictable component in consumption growth can attenuate the high risk aversion implications of standard asset pricing models with recursive utility preferences. However, this persistent component is difficult to detect in the data. Croce et al. (2006) present a limited information economy where agents face a signal extraction problem. Their model addresses the identification issues of the long run risk component. Hansen and Sargent (2007b) is another example for the difficulty in identifying the long run risk component. However, in addition to a signal extraction problem, their agent seeks robust policies and consequently his estimation procedure is modified.

³⁶See table IV in Bansal and Yaron (2004).

is much less persistent, the short rate can potentially be very volatile. If shocks to μ were to die out much slower than shocks to v , the ergodic distribution of the short rate would be very volatile. In that sense, we might be able to reconcile our results with the calibration exercise of Bansal and Yaron (2004) if we assumed an expected consumption growth rate process.

Expressions (4.5) and (4.6) allow us to discuss a mechanism which is central to our results. In Figure 5.5 we plot the objective and perceived impulse response functions for the conditional variance v following a shock. Note that, unlike a rational expectations agent, the robust agent is on average wrong about the future evolution of v . Hence, his biased expectations lead him to believe that the conditional variance will decrease. As mentioned earlier, this should lead to an upward-sloping unconditional sloping yield curve through the precautionary savings channel.

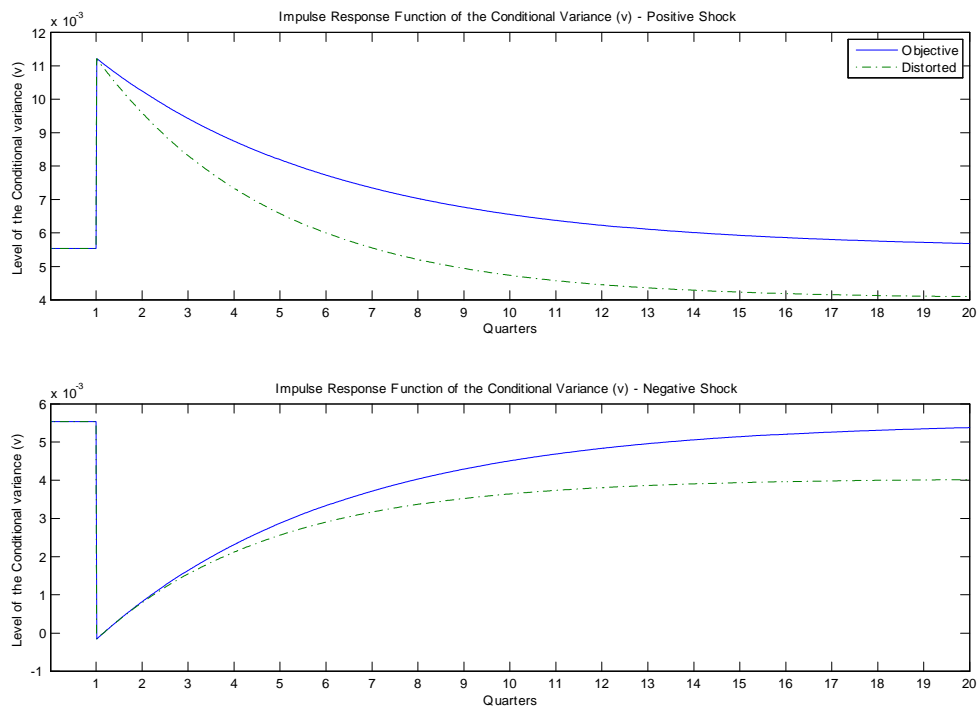


Figure 5.5: Biased expectations. Using the parameters estimated over the entire period $Q2.52 - Q4.06$, the figure shows the impulse response function of the conditional variance to a positive and negative shocks. The solid line represents the objective evolution of v and the dashed line represents what the robust agent believes the evolution of v is going to be.

5.2.3 Theoretical and Empirical Moments

Table 5.5 presents a comparison of model-implied and empirical moments over different time spans for the equity and consumption data. Table 5.6 presents the same exercise, but without imposing consumption growth rate volatility in the estimation. The model fares well, especially with the aggregate market return and the equity premium.³⁷ Also, the model is doing a good job in matching the low consumption growth rate. The same conclusion seems to hold over different time horizons. Note, however, that again we see the tension between market return and consumption growth volatility. When imposing consumption growth volatility, the model compromises on the implied market return volatility being somewhere between the empirical consumption growth rate volatility and the empirical market return volatility. When ignoring the consumption growth volatility from the estimation procedure, the model easily matches the aggregate market return volatility. This result is obviously not surprising since we have a log-agent that consumes a constant fraction of his wealth. Given that the substitution effect and the income effects cancel each other, the agent absorbs all market fluctuations to his marginal utility.

Tables 5.7 and 5.8 report model implied and empirical moments for the bond market, where the second table ignores the volatility of consumption growth in our set of moments. The model is doing a good job in reproducing the levels of the 3-months and 1-year real yields. The results in the last two columns of each table are particularly interesting. The second to last column ($\rho(\mathcal{Y}_{3m})$) reports the first-order autocorrelation of the 3-months yield. Note that we do not impose this restriction in our estimation and yet the model is able to produce this moment with high accuracy. This information is indirectly encoded into the orthogonality conditions through the imposition of the change in the 3-months yield. The last column captures the holding period returns of a strategy that dictates buying a 1-year bond and selling it after 3 quarters. Backus et al. (1989) point to the difficulty of representative agent models to account for both the sign and magnitude of holding period returns in the bond market. Again we note that we did not impose any holding period returns conditions in the estimation procedure and yet the model captures the returns dynamics well. Nevertheless, we should note that by imposing the spread and the change in the short rate conditions, we provide the estimation procedure with enough information about the dynamics of the 1-year and 3-months yields to the extent that the holding period returns are captured accurately by the model.

The top panel in Figure 5.6 presents estimation results over the years '97 – '06. During this period TIPS bonds were traded in the U.S. and thus provide a good proxy to real yields. The

³⁷Erbas and Mirakhor (2007) document global evidence (53 emerging and mature markets) that a large part of the equity premium reflects investor aversion to ambiguities resulting from institutional weaknesses.

Table 5.5: Empirical and theoretical equity and goods market moments (with consumption volatility restriction). The period column represents the time interval of the data that is used to estimate the model. The data is in quarterly frequency with quarterly values. T is the number of quarterly observations used to estimate the model. Columns with the number (1) present the empirical moments. Empirical moments computed with the data and theoretical moments are implied by the estimated model. Columns with the number (2) present the theoretical moments. The theoretical moments were generated using 1,000 replications of the economy that was calibrated using the estimated parameters over the corresponding period. Robust standard errors are given below each moment. The standard errors were corrected using the Newey-West procedure with 4 lags. The standard errors for the theoretical moments were computed over the 1,000 replications. All moments are given in % values. μ_R , μ_C , σ_R , and \mathcal{Y}_{3m} are the real return on the market (including dividends), real growth rate of consumption, volatility of real aggregate market return and real 3 month yield, respectively.

Period	T	μ_R		$\mu_R - \mathcal{Y}_{3m}$		σ_R		μ_C	
		(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
$Q2.52 - Q4.06$	218	12.820	13.692	11.289	11.719	34.306	22.385	2.109	2.924
		2.321	0.024	2.350	0.034	2.164	0.038	0.174	0.022
$Q1.62 - Q4.06$	180	11.879	12.586	10.130	10.552	35.499	22.748	2.085	2.940
		2.544	0.028	2.562	0.039	2.487	0.044	0.188	0.025
$Q1.72 - Q4.06$	140	12.611	12.985	10.925	10.879	36.599	23.302	1.937	2.882
		3.021	0.033	3.029	0.044	2.865	0.045	0.204	0.030
$Q1.82 - Q4.06$	100	15.424	14.614	13.234	12.218	35.510	24.372	2.206	3.259
		3.240	0.039	3.206	0.051	2.941	0.041	0.178	0.035
$Q1.90 - Q4.06$	68	13.060	12.298	11.393	10.579	34.075	23.315	2.035	2.971
		3.887	0.044	3.807	0.057	3.753	0.046	0.203	0.040
$Q2.52 - Q4.81$	118	10.749	17.617	9.784	16.053	32.983	24.751	2.004	2.977
		3.155	0.042	3.276	0.062	3.111	0.097	0.279	0.036
$Q2.52 - Q4.89$	149	12.710	15.950	11.241	13.842	34.411	23.964	2.143	3.086
		2.905	0.034	2.976	0.048	2.642	0.063	0.236	0.029

Table 5.6: Empirical and theoretical equity and goods market moments (without consumption volatility restriction). The period column represents the time interval of the data that is used to estimate the model. The data is in quarterly frequency with quarterly values. T is the number of quarterly observations used to estimate the model. Columns with the number (1) present the empirical moments. Empirical moments computed with the data and theoretical moments are implied by the estimated model. Columns with the number (2) present the theoretical moments. The theoretical moments were generated using 1,000 replications of the economy that was calibrated using the estimated parameters over the corresponding period. Robust standard errors are given below each moment. The standard errors were corrected using the Newey-West procedure with 4 lags. The standard errors for the theoretical moments were computed over the 1,000 replications. All moments are given in % values. μ_R , μ_C , σ_R and \mathcal{Y}_{3m} are the real return on the market (including dividends), real growth rate of consumption, volatility of real aggregate market return and real 3 month yield, respectively.

Period	T	μ_R		$\mu_R - \mathcal{Y}_{3m}$		σ_R		μ_C	
		(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
$Q2.52 - Q4.06$	218	12.820	19.564	11.289	17.682	34.306	36.566	2.109	4.110
		2.321	0.043	2.350	0.056	2.164	0.080	0.174	0.037
$Q1.62 - Q4.06$	180	11.879	18.579	10.130	16.616	35.499	36.742	2.085	4.117
		2.544	0.050	2.562	0.064	2.487	0.095	0.188	0.043
$Q1.72 - Q4.06$	140	12.611	18.648	10.925	16.713	36.599	37.482	1.937	4.139
		3.021	0.057	3.029	0.072	2.865	0.094	0.204	0.049
$Q1.82 - Q4.06$	100	15.424	18.802	13.234	16.548	35.510	37.124	2.206	4.496
		3.240	0.063	3.206	0.079	2.941	0.081	0.178	0.056
$Q1.90 - Q4.06$	68	13.060	15.531	11.393	13.782	34.075	34.177	2.035	3.992
		3.887	0.070	3.807	0.087	3.753	0.087	0.203	0.062
$Q2.52 - Q4.81$	117	10.592	25.701	9.625	24.142	33.078	41.079	2.025	4.257
		3.204	0.083	3.327	0.109	3.132	0.231	0.278	0.067
$Q2.52 - Q4.89$	149	12.710	23.824	11.241	21.822	34.411	39.497	2.143	4.347
		2.905	0.062	2.976	0.083	2.642	0.148	0.236	0.052

Table 5.7: Empirical and theoretical bond market moments (with consumption volatility restriction). The period column represents the time interval of the data that is used to estimate the model. The data is in quarterly frequency with quarterly values. T is the number of quarterly observations used to estimate the model. Columns with the number (1) present the empirical moments. Empirical moments computed with the data and theoretical moments are implied by the estimated model. Columns with the number (2) present the theoretical moments. The theoretical moments were generated using 1,000 replications of the economy that was calibrated using the estimated parameters over the corresponding period. Robust standard errors are given below each moment. The standard errors were corrected using the Newey-West procedure with 4 lags. The standard errors for the theoretical moments were computed over the 1,000 replications. All moments, aside from the autocorrelations, are given in % values. \mathcal{Y}_{3m} , \mathcal{Y}_{1y} , and $\rho(\mathcal{Y}_{3m})$ are the real 3 month yield, real 1 year yield and the first order autocorrelation coefficient of the real 3 month yield, respectively. The last column reports real holding period return for buying a one year to maturity bond and selling it after three quarters.

Period	T	\mathcal{Y}_{3m}		\mathcal{Y}_{1y}		$\rho(\mathcal{Y}_{3m})$		$\ln \left[\frac{p(1;v_{t+3})}{p(4;v_t)} \right]$	
		(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
$Q2.52 - Q4.06$	218	1.531	1.973	2.250	2.659	0.843	0.863	2.465	2.888
		0.263	0.010	0.234	0.007	0.060	0.000	0.274	0.006
$Q1.62 - Q4.06$	180	1.749	2.033	2.241	2.501	0.835	0.881	2.394	2.658
		0.292	0.011	0.279	0.008	0.071	0.000	0.321	0.007
$Q1.72 - Q4.06$	140	1.686	2.106	2.214	2.601	0.830	0.876	2.359	2.766
		0.362	0.012	0.351	0.008	0.075	0.001	0.406	0.007
$Q1.82 - Q4.06$	100	2.190	2.396	2.742	2.926	0.862	0.888	2.932	3.105
		0.366	0.012	0.378	0.009	0.062	0.001	0.443	0.009
$Q1.90 - Q4.06$	68	1.666	1.719	2.015	2.073	0.897	0.892	2.118	2.190
		0.422	0.013	0.371	0.011	0.058	0.001	0.420	0.010
$Q2.52 - Q4.81$	118	0.965	1.565	1.865	2.467	0.804	0.800	1.976	2.768
		0.330	0.020	0.247	0.011	0.096	0.001	0.275	0.008
$Q2.52 - Q4.89$	149	1.469	2.107	2.358	2.934	0.823	0.834	2.586	3.211
		0.328	0.015	0.292	0.009	0.077	0.001	0.348	0.007

Table 5.8: Empirical and theoretical bond market moments (without consumption volatility restriction). The period column represents the time interval of the data that is used to estimate the model. The data is in quarterly frequency with quarterly values. T is the number of quarterly observations used to estimate the model. Columns with the number (1) present the empirical moments. Empirical moments computed with the data and theoretical moments are implied by the estimated model. Columns with the number (2) present the theoretical moments. The theoretical moments were generated using 1,000 replications of the economy that was calibrated using the estimated parameters over the corresponding period. Robust standard errors are given below each moment. The standard errors were corrected using the Newey-West procedure with 4 lags. The standard errors for the theoretical moments were computed over the 1,000 replications. All moments, aside from the autocorrelations, are given in % values. \mathcal{Y}_{3m} , \mathcal{Y}_{1y} , and $\rho(\mathcal{Y}_{3m})$ are the real 3 month yield, real 1 year yield and the first order autocorrelation coefficient of the real 3 month yield, respectively. The last column reports real holding period return for buying a one year to maturity bond and selling it after three quarters.

Period	T	\mathcal{Y}_{3m}		\mathcal{Y}_{1y}		$\rho(\mathcal{Y}_{3m})$		$\ln \left[\frac{p(1;v_{t+3})}{p(4;v_t)} \right]$	
		(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
$Q2.52 - Q4.06$	218	1.531	1.882	2.250	2.525	0.843	0.864	2.465	2.739
		0.263	0.013	0.234	0.009	0.060	0.000	0.274	0.008
$Q1.62 - Q4.06$	180	1.749	1.962	2.241	2.417	0.835	0.874	2.394	2.570
		0.292	0.014	0.279	0.010	0.071	0.001	0.321	0.009
$Q1.72 - Q4.06$	140	1.686	1.935	2.214	2.422	0.830	0.868	2.359	2.585
		0.362	0.015	0.351	0.011	0.075	0.001	0.406	0.010
$Q1.82 - Q4.06$	100	2.190	2.253	2.742	2.777	0.862	0.885	2.932	2.955
		0.366	0.017	0.378	0.013	0.062	0.001	0.443	0.012
$Q1.90 - Q4.06$	68	1.666	1.749	2.015	2.092	0.897	0.891	2.118	2.205
		0.422	0.018	0.371	0.015	0.058	0.001	0.420	0.014
$Q2.52 - Q4.81$	117	0.967	1.559	1.828	2.381	0.806	0.800	1.963	2.650
		0.333	0.027	0.237	0.016	0.096	0.001	0.278	0.013
$Q2.52 - Q4.89$	149	1.469	2.002	2.358	2.798	0.823	0.837	2.586	3.065
		0.328	0.021	0.292	0.014	0.077	0.001	0.348	0.011

solid line is the average level of the yield curve over this period with 95% confidence bands. The dot-dashed line is the model-implied average yield curve. Note that we only impose two bond market restrictions in the estimation procedure and yet the model can closely imitate the behavior of the entire yield curve (within the confidence bands). The bottom panel depicts the term structure of the volatilities of yields. Clearly, the model can replicate the downward slope due to the mean reversion in the estimated conditional variance process, as discussed earlier. The impression is that the procedure anchors the implied first and second moments of the 1-year yield to its empirical counterpart, but it is still doing a good job in approximating the entire curve.

As discussed earlier, our model can reconcile the difficulty one factor models face when trying to match both the high persistence of yields and the high convexity of the curve. Figure 5.6 shows that the agent prices the yield curve as if shocks to v die out fast. However, Tables 5.7 and 5.8 confirm that the model can still match the persistence of the short rate. Empirically, all yields exhibit the same level of persistence.³⁸

5.3 ‘Disciplining Fear’: Detection Error probabilities

In this section, we undertake the task of interpreting θ . We showed so far that the model can account for different asset pricing facts and puzzles. Nevertheless, we have yet to tackle an important question - does the model imply too much uncertainty aversion? Even though we showed that coefficients of relative risk aversion and elasticity of intertemporal substitution of unity are sufficient, we still need to gauge the amount of ambiguity aversion implied by the data. Detection error probabilities (DEP’s) are the mechanism through which we can interpret θ , and consequently, assess the amount of ambiguity aversion implied by our estimation.

In order to quantify ambiguity aversion, we ask the following: when the agent examines the (finite amount of) data available to him and has to decide whether the reference or the distorted model generated the data, what is the probability of making a model detection mistake? If the probability is very low, this indicates that the two models are far apart statistically, and that the agent should easily be able to distinguish between them. In this case, one might be led to conclude that the degree of robustness implied by our estimation is unreasonably high. If to the contrary, the DEP is high, then it is reasonable to believe that the agent would find it difficult to determine which model is the true representation of the economy.³⁹

³⁸The term structure literature usually identifies 3 factors that account well for most of the variation in the yield curve (Litterman and Scheinkman (1991)): level, slope and curvature. The level slope is very persistent and, thus, accounts for most of the observed persistence of yields.

³⁹For an elaborate discussion of DEP’s see, for example, Anderson et al. (2003) and Barillas et al. (2007). For a textbook treatment of robustness and DEP’s see chapters 9 and 10 in Hansen and Sargent (2007a).

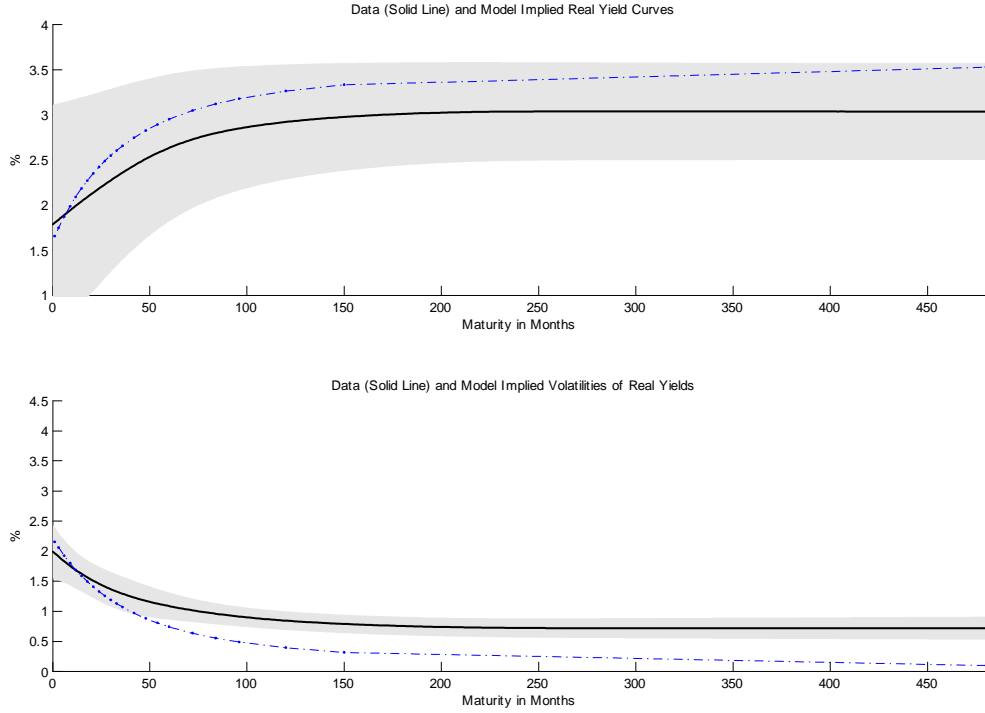


Figure 5.6: Top panel: average real yield curve extracted from the TIPS data from $M1.97 - M12.06$ (solid line) with 95% confidence bands with Newey-West (12 lags) correction. Model implied average yield curve (dot-dashed line). The model is estimated over the same period as the empirical yield curve. Bottom panel: empirical term structure of unconditional volatilities of the TIPS data (solid line). with 95% confidence bands with Newey-West (12 lags) correction. The model is estimated over the same period as the empirical yield curve.

Technically, DEP's are a mapping from the space of structural parameters to a probability space, which is inherently more easily interpretable than parameter values. Based on our estimate of the parameter θ , we infer the detection error probabilities from the data. It then allows us to interpret whether the degree of ambiguity aversion in our parameterization seems excessive. Appendix E details how to derive the DEP for a given economy using simulations.

The last column in tables 5.2 and 5.3 presents the implied DEP's in each economy. First, it is important to point out that DEPs have to be between 0% and 50% (if both models are the same, then there is a 50% chance of making a mistake when assessing which model is the true one). What we find is that our implied DEPs are definitely not unreasonably small, particularly in the context of a framework where the only source of uncertainty is a single shock. This is, once again, an outcome of the interaction between the two main building blocks of our model - robust decision making and state dependent volatility. Together they imply a high enough

market price of risk and uncertainty, and in fact with stochastic volatility the agent does not need to distort the reference model ‘too much’. Therefore, the DEPs are sufficiently large.

The lowest DEP is for our benchmark model. This is not surprising for two reasons. First, we use the longest possible sample, making it easier for the agent/econometrician to distinguish between the objective and distorted models. Second, imposing (the very low) consumption growth rate volatility restricts the implied volatility of the SDF severely. Therefore, the model implies a stronger distortion in a way that enables us to achieve the Hansen-Jagannathan bound. When either the number of observations is smaller or we ignore the consumption growth rate volatility, the DEP increases.

Figure 5.7 present two comparative statics exercises on the implied DEPs. The left panel fixes the benchmark model and varies only θ . The right panel introduces variation only in the number of observations available to the econometrician. We see a clear pattern: Higher θ means less robustness. Thus, it becomes harder to statistically distinguish between the reference and the distorted models as the agent distorts less and less. As $\theta \rightarrow \infty$ the DEP reaches 0.5. This is not surprising, since $\theta = \infty$ implies that the distortion to the DGP is zero (recall (3.20)) and both models are therefore indistinguishable. On the other hand, a lower value of θ means more robustness and the models become statistically distant from each other (in the relative entropy sense), reflected in a lower DEP. Similarly, more observations reduce the DEP, in line with our earlier intuition.

5.3.1 The Evolution of ‘Fear’

In this subsection we document the way fear of model misspecification evolved over time in the context of our framework. We constructed Figure 5.8 by estimating our complete model using rolling (overlapping) windows of 20 years of quarterly data, from the early 1970s to 2007. For any given estimation iteration, we present the point estimate of θ with its corresponding 95% confidence interval and DEP. It is apparent from this figure that θ and DEPs are closely related to each other, with a cross correlation of 0.8113 and Newey-West standard errors with 4 lag correction of 0.0418. Therefore, it strongly confirms the suggestion that we should examine DEPs when trying to understand the level of uncertainty aversion exhibited by economic agents.

On the basis of this exercise it seems that the agent was seeking more robustness in the later period of the sample. An interesting question is to determine whether this evolution could be linked to macroeconomic and financial developments over the same time period, and in particular its link with the discussion about the Great Moderation.⁴⁰ Since the investigation of

⁴⁰On the one hand, macro volatility, and in particular consumption growth volatility, has steadily declined in the later period of our sample (the Great Moderation). However, market return volatility does not exhibit

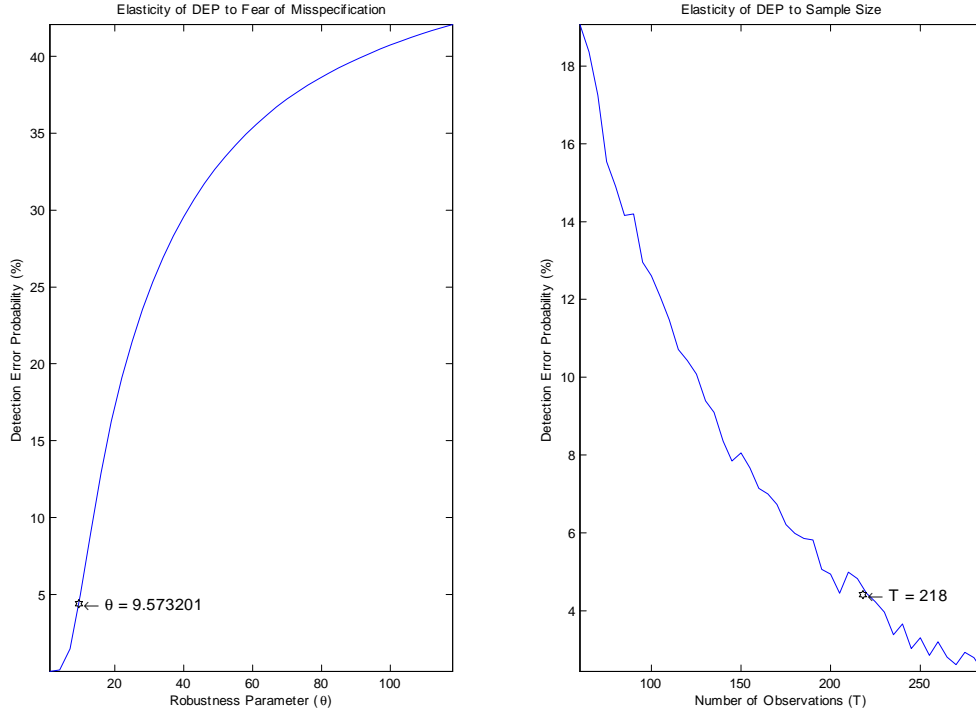


Figure 5.7: Comparative statics on DEP's. Left panel: Fixing all the estimated parameters in the benchmark case with consumption growth volatility as a restriction and over the longest sample $Q2.52 - Q4.06$. Varying robustness in the model (by varying θ on the x-axis) we compute the implied DEP's (y-axis). Right panel: Fixing all the parameters in the benchmark model and varying only the hypothetical number of observations (x-axis) and computing the implied DEP's (y-axis).

any causality is outside the scope of this study, we leave this question for future research.

6 Conclusion

We presented an equilibrium dynamic asset pricing model that can account for key regularities in the market for default free bonds, while predicting an equity premium, risk free rate and consumption growth as in the data. We estimated the model and showed that it performs well, even though the structural parameters of risk aversion and elasticity of intertemporal substitution are unitary. The results are driven by the interaction of the robust control decision

the same pattern. One possibility is that the smoother consumption growth is interpreted by the estimation procedure as an increased uncertainty aversion which implies the decline in detection error probabilities in the later part of the sample. In other words, it is harder to achieve the HJ bound with smoother consumption. Thus, the estimation procedure compensates for this difficulty by encoding more robustness into the agent's behavior. Consequently, the implied DEPs are higher.

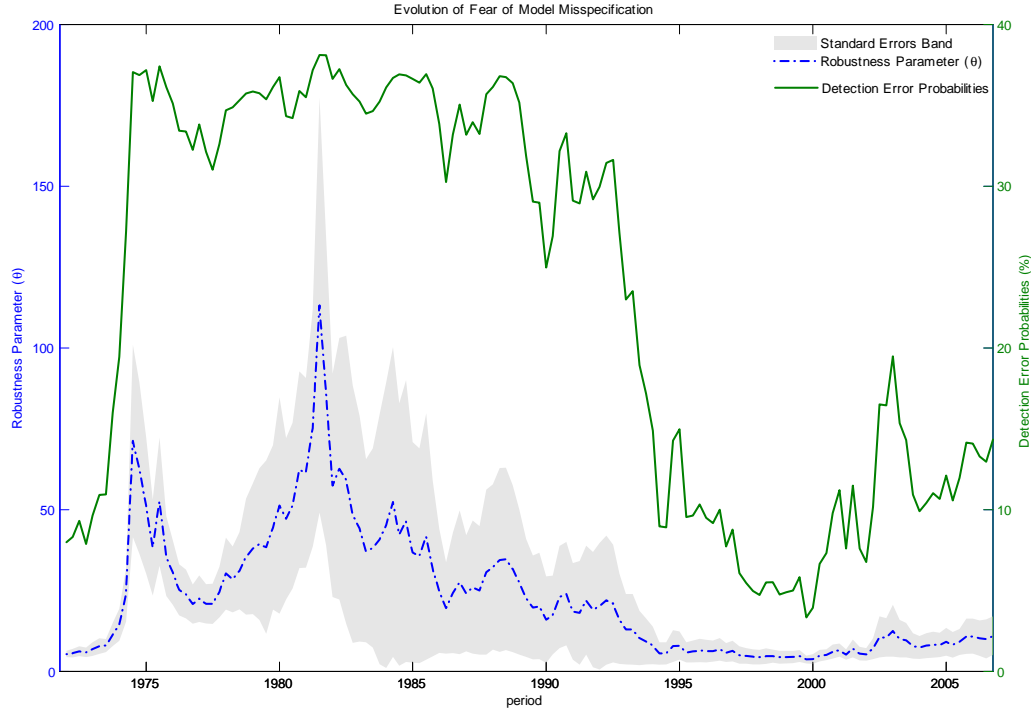


Figure 5.8: Evolution of θ and DEP's over time. We estimate our complete model by using rolling (overlapping) windows of 20 years of quarterly data. For any given estimation iteration, we present the point estimate of θ (dot-dashed line) with its corresponding 95% confidence interval (gray bands). The standard errors are corrected using the Newey-West procedure with 4 lags. Given the estimated parameters in each iteration, we compute the implied DEP (solid line).

mechanism and state dependent conditional volatility of consumption growth. We interpreted most of what is usually considered risk premium as a premium for Knightian uncertainty. The agent is being compensated in equilibrium for bearing the possibility of model misspecification. We also showed that modeling robustness can help explain biases in expectations documented in surveys.

We showed that under the assumption of state dependent conditional volatility of consumption growth, not only the market price of risk is stochastic but also the market price of model uncertainty. As part of our research agenda, we are currently investigating a model with heterogeneous robust control agents. Such a model can generate both state dependent risk and uncertainty premia even though the conditional volatility of consumption is constant. The channel through which the model generates stochastic market prices of risk and uncertainty is the trade between the agents and the consequent fluctuations in the agent's relative wealth.⁴¹

⁴¹Liu et al. (2005) introduce state dependent market price of uncertainty by modeling rare events. Hansen

We also suggested that different frequencies in the conditional volatility of consumption growth are potentially important in understanding asset prices. We find it easier to detect high frequency variation in the volatility of consumption growth rate. Also, the full estimation of the model has trouble detecting the lower frequency component. We believe that further investigation of this point is warranted. In addition, an interesting extension would be to consider the link between the evolution of volatility over time and the behavior of asset prices, in the presence of ambiguity aversion. This is directly linked to the recent literature on the Great Moderation in macroeconomics.

We also believe that extending the empirical investigation to a broader asset class can be fruitful. Liu et al. (2005), for example, examine options data in the context of a robust equilibrium with rare events. We believe that one can address different empirical regularities pertaining to the valuation of interest rate sensitive assets with robust considerations. Also, we think that robustness can shed more light on our understanding of exchange rate dynamics, and in particular the failure of uncovered interest rate parity. Finally, our model is a complete characterization of a real economy. One can extend this framework to a nominal one either by assuming an exogenous price level process as in Cox et al. (1985) and Wachter (2001) or by modelling an exogenous money supply process as in Buraschi and Jiltsov (2005) to derive an endogenous price level.

and Sargent (2007b) introduce state dependent market price of uncertainty through the distortion (tilting) of Bayesian model averaging.

References

- Abel, A., (2002), "An exploration of the effects of pessimism and doubt on asset returns." *Journal of Economic Dynamics & Control* 26, 1075–1092.
- Anderson, E., L. P. Hansen, and T. J. Sargent, (2003), "A quartet of semigroups for model specification, robustness, prices of risk, and model detection." *Journal of the European Economic Association* 1 (1):68–123.
- Backus, D., Foresi, S., Mozumdar, A. and Wu, L., (1998), "Predictable changes in yields and forward rates." NBER WP no 6379.
- Backus, D., Gregory, A. W. and Zin, S. E., (1989), "Risk premiums in the term structure: Evidence from artificial economies." *Journal of Monetary Economics* 24, 371-99.
- Bansal, R., Katchatrian V., and A. Yaron (2005), "Interpretable asset markets?" *European Economic Review* 49, 531-560.
- Bansal, R., Kiku, D., and Yaron, A. (2007), "Risks for the long run: estimation and inference." Working paper, Duke and UPenn.
- Bansal, R., and A. Yaron. (2004), "Risks for the long run: a potential resolution of asset pricing puzzles." *Journal of Finance* 59(4):1481–509.
- Barillas, F., Hansen, L. P., Sargent, T. J., (2007), "Doubts or variability?" Working paper, University of Chicago and NYU.
- Bekaert, G., Hodrick, R. and Marshall, D., (1997), "On biases in tests of the expectations hypothesis of the term structure of interest rates." *Journal of Financial Economics* 44, 309-48.
- Bliss, R. (1999), "Fitting term structures to bond prices." Working paper, Federal Reserve Bank of Chicago.
- Brown, S. J. and Dybvig, P. H. (1986), "The empirical implications of the Cox, Ingersoll, and Ross theory of the term structure of interest rates." *Journal of Finance* 41, 617-32.
- Brown, R. H. and Schaefer, S. (1994), "The term structure of real interest rates and the Cox, Ingersoll, and Ross Model." *Journal of Financial Economics* 35, 3-42.
- Buraschi, A., Jiltsov, A., (2005), "Inflation risk premia and the expectations hypothesis." *Journal of Financial Economics* 75, 429-490.

- Cagetti, M., Hansen, L. P., Sargent, T. and Williams, N. (2002), "Robustness and pricing with uncertain growth." *Review of Financial Studies* 15, 363–404.
- Campbell, J. Y., (2000), "Asset pricing at the millennium." *Journal of Finance* 55(4): 1515–1567.
- Campbell, J. Y. and Shiller, R., (1991), "Yield spreads and interest rates: A bird's eye view." *Review of Economic Studies* 58, 495–514.
- Chan, K, Karolyi, A., Longstaff, F. and Sanders, A., (1992), "The volatility of the short-term interest rates: An empirical comparison of alternative models of the term structure of interest rates." *Journal of Finance* 68: 1209–27.
- Chapman, D. (1997), "The cyclical properties of consumption growth and the real term structure." *Journal of Monetary Economics* 39, 145–72.
- Cochrane, J. H. and Piazzesi, M., (2002), "Bond risk premia." *American Economic Review* 95, 138–160.
- Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985), "A theory of the term structure of interest rates." *Econometrica* 53, 385–407.
- Croce, M. M., Lettau, M., Ludvigson, S. C. (2006), "Investor information, long-run risk, and the duration of risky cash flows." Working paper, NYU.
- Dai, Q. (2003), "Term structure dynamics in a model with stochastic internal habit." Working paper, UNC.
- Dai, Q. and Singleton, K. (2000), "Specification analysis of affine term structure models." *Journal of Finance* 55, 1943–78.
- Dai, Q. and Singleton, K. (2002), "Expectation puzzles, time-varying risk premia, and affine models of the term structure." *Journal of Financial Economics* 63, 415–41.
- Dai, Q. and Singleton, K. (2003), "Term structure modeling in theory and reality." *Review of Financial Studies* 16(3), 631–678.
- Donaldson, J. B., Johnsen, T. H. and Mehra, R. (1990), "On the term structure of interest rates." *Journal of Economic Dynamics and Control* 14, 571–596.
- Drost, F.C., Nijman, T.E. (1993), "Temporal Aggregation of Garch Processes." *Econometrica* 61, 909–927.

- Duffie, D., Epstein, L., (1992a), "Asset pricing with stochastic differential utility." *Review of Financial Studies* 5, 411–436.
- Duffie, D., and Epstein, L., (1992b), "Stochastic differential utility." *Econometrica* LX, 353–394.
- Duffie, D. and Kahn, R. (1996), "A yield-factor model of interest rates." *Mathematical Finance* 6, 379–406.
- Duffie, D., and Singleton, K., (1993), "Simulated methods estimation of markov models of asset prices." *Econometrica* 61, 929–952.
- Ellsberg, D. (1961), "Risk, ambiguity and the savage axioms." *Quarterly Journal of Economics* 25, 643–69.
- Epstein, L. G. and Wang, T., (1994), "Intertemporal asset pricing under knightian uncertainty." *Econometrica* 62(2):283–322.
- Epstein, L. G. and Schneider, M., (2003), "Recursive multiple priors." *Journal of Economic Theory* 113:32–50.
- Epstein, L. and Zin, S., (1989), "Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework." *Econometrica* 57, 937–969.
- Erbaş, S. N. and Mirakhor, A. (2007), "The equity premium puzzle, ambiguity aversion, and institutional quality." IMF WP/07/230.
- Fama, E. F. and Bliss, R.R. (1987), "The information in long-maturity forward rates." *American Economic Review* 77, 680–92.
- Gagliardini, P., Porchia, P. and Trojani, F., (2004), "Ambiguity aversion and the term structure of interest rates." Working paper, University of St.Gallen.
- Gibbons, M. R. and Ramaswamy, K., (1993), "A test of the Cox, Ingersoll and Ross model of the term structure." *Review of Financial Studies* 6, 619–58.
- Gilboa, I. and Schmeidler, D., (1989), "Maxmin expected utility with non-unique prior." *Journal of Mathematical Economics* 18:141–153.
- Giordani, P., Söderlind, P., (2006), "Is there evidence of pessimism and doubt in subjective distributions? Implications for the equity premium puzzle." *Journal of Economic Dynamics & Control* 30, 1027–1043.

- Glosten, L., Jagannathan, R., Runkle, D. (1993), "On the relation between the expected value and the volatility of the nominal excess return on stocks." *Journal of Finance* 48, 1779-1801.
- Hansen, L. P., (1982), "Large sample properties of generalized method of moments estimators." *Econometrica* 50, 1029-54.
- Hansen, L. P. and Jagannathan, R., (1991), "Implications of security market data for models of dynamic economies." *Journal of Political Economy* 99:225-262.
- Hansen, L. P. and Sargent, T. J., (2007a), "Robustness." Princeton, New Jersey: Princeton University Press.
- Hansen, L.P. and Sargent, T. J., (2007b), "Fragile beliefs and the price of model uncertainty." Working paper, NYU and Univ. of Chicago.
- Hansen, L. P., Sargent, T. J., Turmuhambetova, G. A. and Williams, N. (2006), "Robust control, min-max expected utility, and model misspecification." *Journal of Economic Theory* 128, 45-90.
- Hansen, L. P., Sargent, T. and Tallarini, T. (1999), "Robust permanent income and pricing." *Review of Economic Studies* 66, 873-907.
- Litterman, R. and Scheinkman, J. (1991), "Common factors affecting bond returns." *Journal of Fixed Income* 1, 54-61.
- Liu, J., Pan, J. and Wang, T. (2005), "An equilibrium model of rare-event premia and its implication for option smirks." *Review of Financial Studies* 18(1), 131-64.
- Longstaff, F. and Schwartz, E., (1992), "Interest rate volatility and the term structure: A two-factor general equilibrium model." *Journal of Finance* 47, 1259-82.
- Ludwig, A. and Zimmer, A. (2006), "Rational expectations and ambiguity: A comment on Abel (2002)." *Economics Bulletin*, Vol. 4, No. 2 pp. 1-15.
- Maenhout, P. (2004), "Robust portfolio rules and asset pricing." *Review of Financial Studies* 17, 951-983.
- Maccheroni, F., Marinacchi, M., Rustichini, A. (2006) "Ambiguity aversion, robustness, and the variational representation of preferences." *Econometrica* 74, 1447-1498.

- McCulloch, J. H. and Kwon, H. C. (1993), "U.S. term structure data, 1947-1991." Working paper, Ohio State University.
- Mehra, R. and Prescott, E., (1985), "The equity premium: a puzzle." *Journal of Monetary Economics* 15, 145–61.
- Miao, J., (2004), "A note on consumption and savings under knightian uncertainty." *Annals of Economics and Finance* 5, 299-301.
- Nelson, D., (1991), "Conditional heteroskedasticity in asset returns: A new approach." *Econometrica* 59, 347-370.
- Strzalecki, T. (2007), "Axiomatic foundations of multiplier preferences." Working paper, Northwestern University.
- Piazzesi, M. (2004), "Affine term structure models", Y. Aït-Sahalia and L. P. Hansen (eds.), *Handbook of Financial Econometrics*, Amsterdam: North Holland.
- Routledge, B. R., and Zin, S. E. (2001), "Model uncertainty and liquidity." Working paper, NBER.
- Skiadas, C. (2003), "Robust control and recursive utility." *Finance and Stochastics* 7, 475–489.
- Söderlind, P., (2006), "C-CAPM without ex-post data." WP 344, *NCCR FINRISK*
- Stock, J. H. and Watson, M. W., (2003), "Has the business cycles changed? evidence and explanations." in "Monetary policy and uncertainty: adapting to a changing economy." pp. 9-56. Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming.
- Uppal, R., Wang, T., (2003), "Model misspecification and underdiversification." *Journal of Finance* 58, 2465–2486.
- Wachter, J., (2006), "A consumption-based model of the term structure of interest rates." *Journal of Financial Economics* 79, 365–399.
- Weil, P., (1989), "The equity premium puzzle and the risk free rate puzzle." *Journal of Monetary Economics* 24, 401–421.
- Weil, P., (1990), "Nonexpected Utility in Macroeconomics." *Quarterly Journal of Economics* 105 (1):29–42.
- Whitelaw, R., (1994), "Time variation and covariations in the expectation and volatility of stock market returns." *Journal of Finance* 49, 515–541.

A Proof of Lemma 1

We will prove the lemma in the context of the two period example. Given the class of Normal distributions, we need to calculate the relative entropy between the reference measure \mathbb{P} and an arbitrary Normal distribution $\mathbb{Q} \sim N(\mu_{\mathbb{Q}}, \sigma_{\mathbb{Q}}^2)$. Recall that the relative entropy between two distributions is defined as

$$\mathcal{R}(\mathbb{Q}) \equiv \int \ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) d\mathbb{Q}.$$

We first calculate the integrand

$$\ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) = \ln \frac{\sigma_{\mathbb{P}}}{\sigma_{\mathbb{Q}}} + \left[-\frac{(x - \mu_{\mathbb{Q}})^2}{2\sigma_{\mathbb{Q}}^2} + \frac{(x - \mu_{\mathbb{P}})^2}{2\sigma_{\mathbb{P}}^2} \right].$$

Then, we take expectations with respect to \mathbb{Q}

$$\begin{aligned} \int \ln \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right) d\mathbb{Q} &= \ln \frac{\sigma_{\mathbb{P}}}{\sigma_{\mathbb{Q}}} + \mathbb{E}^{\mathbb{Q}} \left[-\frac{(x - \mu_{\mathbb{Q}})^2}{2\sigma_{\mathbb{Q}}^2} + \frac{(x - \mu_{\mathbb{P}})^2}{2\sigma_{\mathbb{P}}^2} \right] \\ &= \ln \frac{\sigma_{\mathbb{P}}}{\sigma_{\mathbb{Q}}} - \frac{1}{2} + \mathbb{E}^{\mathbb{Q}} \left[\frac{[x - \mu_{\mathbb{P}} + (\mu_{\mathbb{Q}} - \mu_{\mathbb{P}})]^2}{2\sigma_{\mathbb{P}}^2} \right] \\ &= \ln \frac{\sigma_{\mathbb{P}}}{\sigma_{\mathbb{Q}}} - \frac{1}{2} + \frac{\sigma_{\mathbb{Q}}^2 + (\mu_{\mathbb{Q}} - \mu_{\mathbb{P}})^2}{2\sigma_{\mathbb{P}}^2}. \end{aligned}$$

In the log-utility case, we need to minimize (2.1) over the (potentially) distorted mean and variance. First, we take the first order condition with respect to $\mu_{\mathbb{Q}}$

$$\begin{aligned} \underbrace{1}_{\text{Benefit}} + \underbrace{\theta \frac{\mu_{\mathbb{Q}} - \mu_{\mathbb{P}}}{\sigma_{\mathbb{P}}^2}}_{\text{Cost}} &= 0 \\ \implies \mu_{\mathbb{Q}} &= \mu_{\mathbb{P}} - \frac{\sigma_{\mathbb{P}}^2}{\theta}. \end{aligned}$$

Therefore, the mean distortion is additive and equals to $-\frac{\sigma_{\mathbb{P}}^2}{\theta}$. Also, the first order condition with respect to the variance distortion reveals that the robust control agent chooses not to distort the variance

$$\begin{aligned} \underbrace{\theta [-\sigma_{\mathbb{Q}}^{-1} + \sigma_{\mathbb{Q}} \sigma_{\mathbb{P}}^{-2}]}_{\text{Cost}} &= 0 \\ \implies \sigma_{\mathbb{Q}} &= \sigma_{\mathbb{P}} \end{aligned}$$

We next show that when the agent has power utility with risk aversion coefficient $\gamma \neq 1$ he chooses to distort both the mean and variance of the reference distribution. However, the variance distortion has a particular structure. We let $u(C) = C^{1-\gamma}/(1-\gamma)$, $\gamma \neq 1$ and take first order condition with respect to the distorted mean

$$\underbrace{\alpha^{1-\gamma} \exp \left[(1-\gamma) \mu_{\mathbb{Q}} + (1-\gamma)^2 \sigma_{\mathbb{Q}}^2 / 2 \right]}_{\text{Benefit}} + \underbrace{\theta \frac{\mu_{\mathbb{Q}} - \mu_{\mathbb{P}}}{\sigma_{\mathbb{P}}^2}}_{\text{Cost}} = 0. \quad (\text{A.1})$$

And the first order condition with respect to the variance distortion is given by

$$\underbrace{\alpha^{1-\gamma} (1-\gamma) \sigma_{\mathbb{Q}} \exp \left[(1-\gamma) \mu_{\mathbb{Q}} + (1-\gamma)^2 \sigma_{\mathbb{Q}}^2 / 2 \right]}_{\text{Benefit}} + \underbrace{\theta \left[-\sigma_{\mathbb{Q}}^{-1} + \sigma_{\mathbb{Q}} \sigma_{\mathbb{P}}^{-2} \right]}_{\text{Cost}} = 0. \quad (\text{A.2})$$

These two equations can be solved numerically to obtain the optimal distortions. However, we can show that the variance distortion is linked to the mean distortion in a particular way. Divide (A.2) by (A.1) and rearrange to isolate for the distorted variance

$$\sigma_{\mathbb{Q}}^2 = \frac{\sigma_{\mathbb{P}}^2}{1 - (\gamma - 1) (\mu_{\mathbb{P}} - \mu_{\mathbb{Q}})}.$$

Assuming $\gamma > 1$ and the distorted mean $\mu_{\mathbb{Q}} < \mu_{\mathbb{P}}$ then $\sigma_{\mathbb{Q}}^2 > \sigma_{\mathbb{P}}^2$ and $\partial \sigma_{\mathbb{Q}}^2 / \partial (\mu_{\mathbb{P}} - \mu_{\mathbb{Q}}) > 0$. This completes the proof.

B Optimal Policies and Equilibrium

In this appendix we provide some additional details on the derivation of the optimal policies of the agent and the solution of the value function in equilibrium. With a slight abuse of notation, we write the HJB equation as

$$0 = \left[\log C_t + \frac{\theta}{2} h_t^2 \right] dt + \mathbb{E}_t^{\mathbb{Q}} dJ - \rho J dt. \quad (\text{B.1})$$

We posit the following guess for the agent's value function

$$J(W_t, v_t) = \frac{\log W_t}{\rho} + \delta_0 + \delta_1 v_t. \quad (\text{B.2})$$

Applying Ito's lemma to (B.2) and omitting time subscripts for convenience we get

$$\begin{aligned}
dJ &= J_W dW + J_v dv + \frac{1}{2} J_{WW} [dW]^2 \\
&= \left[J_W \mu_W^{\mathbb{Q}} + J_v (a_0 + a_1 v + \sigma_v h \sqrt{v}) + \frac{1}{2} J_{WW} \sigma_W^2 \right] dt + [J_W \sigma_W + J_v \sqrt{v} \sigma_v] dB^{\mathbb{Q}} \\
&= \left[\frac{1}{\rho W} (rW + \alpha W (\mu_R - r) - C + \alpha W \sigma_R h) + \delta_1 (a_0 + a_1 v + \sigma_v h \sqrt{v}) - \frac{\alpha^2 \sigma_R^2}{2\rho} \right] dt \\
&\quad + \left[\frac{\alpha \sigma_R}{\rho} + \delta_1 \sqrt{v} \sigma_v \right] dB^{\mathbb{Q}}.
\end{aligned} \tag{B.3}$$

Next, for the minimization problem we take the derivative of the value function with respect to h and obtain

$$\begin{aligned}
\underbrace{\theta h}_{\text{Cost}} + \underbrace{\frac{\alpha \sigma_R}{\rho} + \delta_1 \sigma_v \sqrt{v}}_{\text{Benefit}} &= 0 \\
\implies h &= -\frac{1}{\theta} \left(\frac{\alpha \sigma_R}{\rho} + \delta_1 \sigma_v \sqrt{v} \right).
\end{aligned}$$

Recall that in equilibrium the agent holds the entire claim on the output process, and thus $\alpha = 1$. This yields expression (3.20).

Deriving α requires taking first order conditions in the maximization problem which shows up only in the drift of dJ . Also, deriving the consumption policy yields the usual envelope type condition $u'(C) = J_W$.

We now solve for the parameters δ_0 and δ_1 . First, plug in (B.2) and (B.3) into (B.1) and use the optimal policies for h , α and C and the equilibrium risk free rate and market return

$$\begin{aligned}
0 &= \log \rho + \log W + \frac{v}{2\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right)^2 - \rho \left(\frac{1}{\rho} \log W + \delta_0 + \delta_1 v \right) \\
&\quad + \frac{1}{\rho W} (rW + \alpha W (\mu_R - r) - C + \alpha W \sigma_R h) + \delta_1 (a_0 + a_1 v + \sigma_v \sqrt{v} h) - \frac{\alpha^2 \sigma_R^2}{2\rho}, \\
0 &= \log \rho + \frac{v}{2\theta} \left(\frac{1}{\rho^2} + \frac{2\delta_1 \sigma_v}{\rho} + \delta_1^2 \sigma_v^2 \right) - \rho \delta_0 - \rho \delta_1 v \\
&\quad + \frac{1}{\rho} \left[\mu - \frac{v}{\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right) \right] + \delta_1 \left(a_0 + a_1 v - \sigma_v \frac{v}{\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right) \right) - \frac{v}{2\rho}.
\end{aligned}$$

Collecting coefficients for v

$$\frac{1}{2\theta} \left(\frac{1}{\rho^2} + \frac{2\delta_1 \sigma_v}{\rho} + \delta_1^2 \sigma_v^2 \right) - \rho \delta_1 - \frac{1}{\rho \theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right) + \delta_1 a_1 - \delta_1 \frac{\sigma_v}{\theta} \left(\frac{1}{\rho} + \delta_1 \sigma_v \right) - \frac{1}{2\rho} = 0.$$

Rearranging to get a quadratic in δ_1

$$\begin{aligned} \left(\frac{\sigma_v^2}{2\theta} - \frac{\sigma_v^2}{\theta}\right) \delta_1^2 + \left(\frac{\sigma_v}{\theta\rho} + a_1 - \rho - \frac{\sigma_v}{\theta\rho} - \frac{\sigma_v}{\theta\rho}\right) \delta_1 + \left(\frac{1}{2\theta\rho^2} - \frac{1}{2\rho} - \frac{1}{\theta\rho^2}\right) &= 0, \\ \frac{\sigma_v^2}{2\theta} \delta_1^2 - \left(a_1 - \rho - \frac{\sigma_v}{\theta\rho}\right) \delta_1 + \left(\frac{1}{2\rho} + \frac{1}{2\theta\rho^2}\right) &= 0. \end{aligned}$$

Solve the quadratic equation

$$\begin{aligned} A &= \frac{\sigma_v^2}{2\theta}, \\ B &= -\left(a_1 - \rho - \frac{\sigma_v}{\theta\rho}\right), \\ C &= \frac{1}{2\rho} + \frac{1}{2\theta\rho^2}. \end{aligned}$$

To prove that it is indeed an equilibrium check that

$$\begin{aligned} B^2 - 4AC &= \left(a_1 - \rho - \frac{\sigma_v}{\theta\rho}\right)^2 - 4\frac{\sigma_v^2}{2\theta} \left(\frac{1}{2\rho} + \frac{1}{2\theta\rho^2}\right) \\ &= \left(a_1 - \rho - \frac{\sigma_v}{\theta\rho}\right)^2 - \frac{\sigma_v^2}{\rho\theta} \left(1 + \frac{1}{\theta\rho}\right) \\ &= (a_1 - \rho)^2 - 2(a_1 - \rho) \frac{\sigma_v}{\theta\rho} - \frac{\sigma_v^2}{\theta\rho} \\ &\geq (a_1 - \rho)^2 - 2(a_1 - \rho) \frac{\sigma_v}{\theta\rho} - \frac{\sigma_v^2}{\theta^2\rho^2} \\ &= \left(a_1 - \rho - \frac{\sigma_v}{\theta\rho}\right)^2 \\ &\geq 0. \end{aligned}$$

The first inequality follows from the fact that $\frac{\sigma_v^2}{\theta\rho} \geq 0$ and $0 \leq \theta\rho < 1$.

With a solution for δ_1 , we find δ_0 by collecting the constant terms

$$\begin{aligned} \log \rho - \rho\delta_0 + \frac{\mu}{\rho} + \delta_1 a_0 &= 0 \\ \implies \delta_0 &= \frac{1}{\rho} \left(\log \rho + \frac{\mu}{\rho} + \delta_1 a_0 \right). \end{aligned}$$

C Pricing the Term Structure

In this appendix we give a more detailed derivation of the bond price. We use the partial differential equation approach which is very common when pricing fixed income securities.

Applying Ito's lemma to (4.2) we derive the dynamics of the bond price with arbitrary maturity

$$\frac{dp_t}{p_t} = -[\beta'_0(\tau) + \beta'_1(\tau)v_t]dt + \beta_1(\tau)dv_t + \frac{1}{2}\beta_1^2(\tau)[dv_t]^2. \quad (\text{C.1})$$

Next, plug (3.2), (3.21), (4.1) and (C.1) into (4.3) to get

$$\begin{aligned} 0 = & -[\beta'_0(\tau) + \beta'_1(\tau)v_t] + \beta_1(\tau)[a_0 + (a_1 + \sigma_v(1 - \phi))v_t] + \frac{1}{2}\beta_1^2(\tau)\sigma_v^2v_t \\ & - (\rho + b_0) + \phi v_t - \beta_1(\tau)\sigma_v v_t. \end{aligned}$$

Collecting the coefficients of v and the free coefficients we get two simple ordinary differential equations. The first is a Riccati equation with constant coefficients

$$\beta'_1(\tau) = \frac{1}{2}\sigma_v^2\beta_1^2(\tau) + (a_1 - \phi\sigma_v)\beta_1(\tau) + \phi,$$

and the second becomes trivial after we solve for β_1

$$\beta'_0(\tau) = \beta_1(\tau)a_0 - (\rho + b_0),$$

with the boundary conditions

$$\beta_0(0) = \beta_1(0) = 0.$$

let $\bar{\beta}_1$ be a particular (constant) solution. In that case $\bar{\beta}'_1 = 0$, and $\bar{\beta}_1$ is given by

$$\bar{\beta}_1 = \frac{-(a_1 - \phi\sigma_v) - \sqrt{(a_1 - \phi\sigma_v)^2 - 2\sigma_v^2\phi}}{\sigma_v^2}.$$

Let $\beta_1 = \bar{\beta}_1 + \frac{1}{z}$. Then

$$\begin{aligned} \left[\bar{\beta}_1(\tau) + \frac{1}{z(\tau)}\right]' &= \frac{1}{2}\sigma_v^2\left[\bar{\beta}_1 + \frac{1}{z(\tau)}\right]^2 + (a_1 - \phi\sigma_v)\left[\bar{\beta}_1 + \frac{1}{z(\tau)}\right] + \phi, \\ -\frac{z'(\tau)}{z^2(\tau)} &= \frac{1}{2}\sigma_v^2\left[\frac{2\bar{\beta}_1}{z(\tau)} + \frac{1}{z^2(\tau)}\right] + (a_1 - \phi\sigma_v)\frac{1}{z(\tau)} \\ \implies z'(\tau) + [\sigma_v^2\bar{\beta}_1 + (a_1 - \phi\sigma_v)]z(\tau) + \frac{1}{2}\sigma_v^2 &= 0. \end{aligned}$$

The solution is derived by simple integration. The boundary condition on z is determined through the boundary condition on β_1 . Since $\beta_1(0) = 0$ we have that $\bar{\beta}_1 + \frac{1}{z(0)} = 0 \implies z(0) = -\frac{1}{\bar{\beta}_1}$. Define

$$\begin{aligned}\Xi &\equiv \sigma_v^2 \bar{\beta}_1 + (a_1 - \phi \sigma_v) \\ &= -\sqrt{(a_1 - \phi \sigma_v)^2 - 2\sigma_v^2 \phi}.\end{aligned}$$

Then,

$$\begin{aligned}z(\tau) e^{\Xi \tau} &= -\frac{1}{2} \sigma_v^2 \int e^{\Xi s} ds + \text{const.} \\ &= -\frac{1}{2} \sigma_v^2 \left[\frac{e^{\Xi s}}{\Xi} \right]_0^\tau + \text{const.} \\ &= -\frac{1}{2} \sigma_v^2 \left[\frac{e^{\Xi \tau} - 1}{\Xi} \right] + \text{const.}\end{aligned}$$

Taking into account the normalizing constant, we get

$$\begin{aligned}z(\tau) &= -\frac{1}{2} \sigma_v^2 \left[\frac{1 - e^{-\Xi \tau}}{\Xi} \right] - \frac{e^{-\Xi \tau}}{\bar{\beta}_1} \\ &= -\frac{\sigma_v^2}{2\Xi} + \left[\frac{\sigma_v^2}{2\Xi} - \frac{1}{\bar{\beta}_1} \right] e^{-\Xi \tau} \\ &= \zeta_0 + \zeta_1 e^{-\Xi \tau}.\end{aligned}$$

Finally, we need to back-out $\beta_0(\tau)$ with the boundary condition $\beta_0(0) = 0$

$$\begin{aligned}\beta_0(\tau) &= a_0 \int \beta_1(s) ds - (\rho + b_0) \tau + \text{const.} \\ &= a_0 \int \left[\bar{\beta}_1 + \frac{1}{z(s)} \right] ds - (\rho + b_0) \tau + \text{const.} \\ &= a_0 \int \frac{ds}{z(s)} - (\rho + b_0 - a_0 \bar{\beta}_1) \tau + \text{const.} \\ &= a_0 \left[\frac{s}{\zeta_0} + \frac{1}{\Xi \zeta_0} \ln \left| \zeta_0 + \zeta_1 e^{-\Xi s} \right| \right]_0^\tau - (\rho + b_0 - a_0 \bar{\beta}_1) \tau + \text{const.} \\ &= a_0 \left[\frac{\tau}{\zeta_0} + \frac{1}{\Xi \zeta_0} \ln \left| \frac{\zeta_0 + \zeta_1 e^{-\Xi \tau}}{\zeta_0 + \zeta_1} \right| \right] - (\rho + b_0 - a_0 \bar{\beta}_1) \tau + \text{const.} \\ &= \frac{a_0}{\Xi \zeta_0} \ln \left| \frac{\zeta_0 + \zeta_1 e^{-\Xi \tau}}{\zeta_0 + \zeta_1} \right| - \left(\rho + b_0 - a_0 \bar{\beta}_1 - \frac{a_0}{\zeta_0} \right) \tau.\end{aligned}$$

Given the bond pricing rule, one can easily price the forward yield curve. Let $f(\tau; v_t)$ be the *instantaneous* forward rate contracted at time t for delivery at time $t + \tau$ (i.e., instantaneous borrowing or lending at time $t + \tau$). Then,

$$\begin{aligned} f(\tau; v_t) &= -\frac{p_\tau(\tau; v_t)}{p(\tau; v_t)} \\ &= -[\beta'_0(\tau) + \beta'_1(\tau) v_t], \end{aligned}$$

where p_τ is the derivative of p with respect to maturity τ .

Similarly, given the prices of all default-free zero-coupon bonds, we can price any arbitrary forward contract. Let $F(\tau, s; v_t)$ be the forward rate (price) contracted at time t for delivery at time $t + \tau$ with maturity $t + s$, where $s \geq \tau$. Then,

$$\begin{aligned} F(\tau, s; v_t) &\equiv \frac{\ln p(\tau; v_t) - \ln p(s; v_t)}{s - \tau} \\ &= \frac{1}{s - \tau} \times [-\tau \mathcal{Y}(\tau; v_t) + s \mathcal{Y}(s; v_t)] \\ &= \frac{1}{s - \tau} \{[\beta_0(\tau) - \beta_0(s)] + [\beta_1(\tau) - \beta_1(s)] v_t\}. \end{aligned}$$

Note that $\lim_{s \downarrow \tau} F(\tau, s; v_t) = f(\tau; v_t)$ and $\lim_{\tau, s \uparrow \infty} F(\tau, s; v_t) = r_t$.

Using forward rates, one can conduct regression analysis as in Fama and Bliss (1987) and Backus et al. (1998) to verify the failure of the expectation hypothesis (return predictability).

D Data

Unless otherwise stated, all data are quarterly from $Q2.1952 - Q4.2006$.

- McCulloch-Kwon-Bliss data set: nominal prices and yields of zero coupon bonds - see McCulloch and Kwon (1993) and Bliss (1999). In the estimation exercises we use only the 3 month and 1 year nominal yields at the quarterly frequency to create the real counterparts. The data we use spans the period $Q2.52 - Q4.96$
- Treasury Inflation-Protected Securities (TIPS) data from McCulloch: real yields from $M1.97 - M12.06$. Although the data is available at higher frequencies, we use only observations at the quarterly frequency
- Quarterly market index (NYSE/AMEX/NASDAQ) including distributions from CRSP

- Quarterly CPI (all items), SA, from the BLS (see FREDII data source maintained by the federal reserve bank of St. Louis for full description)
- Semi-annual inflation expectations from the Livingston survey (maintained by the federal reserve bank of Philadelphia) - period $H1.52 - H1.81$. From $Q3.81$ quarterly inflation expectations data from the Survey of Professional Forecasters (SPF) becomes available
- Quarterly inflation expectations from the SPF maintained by the federal reserve bank of Philadelphia. The sample period covers $Q3.81 - Q4.06$
- Quarterly real Personal Consumption Expenditures (PCE): *services* and *nondurables* from the BEA, SA
- Quarterly real Personal Consumption Expenditures PCE: *imputed services of durables* from the Federal Board of Governors
- Civilian Noninstitutional Population series from the BLS
- Monthly real dividends obtained from Robert Shiller's website over the period $M1.52 - M12.06$ (<http://www.econ.yale.edu/~shiller/data.htm>). This data set was used in the GARCH-GJR exercise

Since we use only real data in the estimation, we convert nominal prices to real ones using the price level data. For the short rate (3 months) we use a 3 year moving average of realized inflation to construct a 3 month ahead expected inflation measure. For the 1 year yield we use both the Livingston and SPF survey data to construct a quarterly series of expected inflation. The SPF is sampled at quarterly frequency but it is available only in the latter part of the sample. We interpolate the semi-annual Livingston data to construct quarterly data using piecewise cubic Hermite interpolation.

E Computing Detection Error Probabilities

In this appendix we shortly discuss how we compute DEP's. The discussion is based on chapter 9 in Hansen and Sargent (2007a). The econometrician observes $\left\{ \frac{\Delta C_{t+1}}{C_t} \right\}_{t=1}^T$ and constructs the log-likelihood ratio of the distorted model relative to the objective model

$$\ell^T = \sum_{t=1}^T \log \frac{f\left(\frac{\Delta C_{t+1}}{C_t} | \theta < \infty\right)}{f\left(\frac{\Delta C_{t+1}}{C_t} | \theta = \infty\right)}.$$

The distorted model is denoted with $f(\cdot|\theta < \infty)$ and the reference model is denoted with $f(\cdot|\theta = \infty)$. The distorted model is selected when $\ell^T > 0$ and the objective model is selected otherwise.

There are two types of detection errors:

1. Choosing the distorted model when actually the reference model generated the data

$$\mathbb{P}(\ell^T > 0|\theta = \infty) = \mathbb{E}(\mathbf{1}_{\{\ell^T > 0\}}|\theta = \infty).$$

2. Choosing the reference model when actually the distorted model generated the data

$$\begin{aligned}\mathbb{P}(\ell^T < 0|\theta < \infty) &= \mathbb{E}(\mathbf{1}_{\{\ell^T < 0\}}|\theta < \infty) \\ &= \mathbb{E}(\exp(\ell^T) \mathbf{1}_{\{\ell^T < 0\}}|\theta = \infty).\end{aligned}$$

Therefore, the *average error* (denoted \wp) with a prior of equiprobable models is

$$\begin{aligned}\wp &= \frac{1}{2} [\mathbb{P}(\ell^T > 0|\theta = \infty) + \mathbb{P}(\ell^T < 0|\theta < \infty)] \\ &= \frac{1}{2} \mathbb{E} \{ \min [\exp(\ell^T), 1] | \theta = \infty \}.\end{aligned}\tag{E.1}$$

We can write an (approximate) transition likelihood ratio as

$$\begin{aligned}\frac{f\left(\frac{\Delta C_{t+1}}{C_t}|\theta < \infty\right)}{f\left(\frac{\Delta C_{t+1}}{C_t}|\theta = \infty\right)} &= \exp \left[-\frac{1}{2} \times \frac{\left(\frac{\Delta C_{t+1}}{C_t} - \mu - h_t \sqrt{v_t}\right)^2 - \left(\frac{\Delta C_{t+1}}{C_t} - \mu\right)^2}{v_t} \right] \\ &= \exp \left[-\frac{1}{2} \times \frac{-2 \left(\frac{\Delta C_{t+1}}{C_t} - \mu\right) (1 - \phi) v_t + (1 - \phi)^2 v_t^2}{v_t} \right] \\ &= \exp \left[\left(\frac{\Delta C_{t+1}}{C_t} - \mu\right) (1 - \phi) - \frac{(1 - \phi)^2 v_t}{2} \right].\end{aligned}$$

We simulate the economy 5,000 times using the point estimates of the parameters and construct a likelihood ratio for each economy. Using (E.1) we can immediately derive \wp .