Doubts and Variability

Rhys Bidder[*]
Federal Reserve Bank of San Francisco

Matthew E. Smith[†]
Federal Reserve Board

First draft: September 4, 2010
This draft: August 4, 2011
Comments Welcome

[*]Federal Reserve Bank of San Francisco; Email: rhys.bidder@sf.frb.org

[†]Federal Reserve Board; Email: matthew.e.smith@frb.gov The authors would like to thank Tom Sargent, David Backus, Francisco Barillas, Tim Cogley, Mark Gertler, Lars Hansen, Anastasios Karantounias, Sydney Ludvigson, Jesse Perla, Chris Tonetti, Håkon Tretvoll, for comments on previous drafts as well as seminar participants at Duke University, The Federal Reserve Bank of Chicago, The Federal Reserve Board, NYU, and Yeshiva University. The views expressed in this paper, and all errors and omissions, should be regarded as those solely of the authors, and are not necessarily those of the individuals listed above, the Federal Reserve Bank of San Francisco, the Federal Reserve Board of Governors or the Federal Reserve System.

Doubts and Variability

Abstract

We examine the asset pricing properties of an endowment economy featuring stochastic volatility and an agent who fears his model is misspecified. Due to the nonlinearities inherent in our stochastic volatility model, we are forced to expand the toolkit of the robust control literature. We propose novel algorithms to characterize and simulate the robust agent's worst case model. Using US consumption data, we estimate the parameters of the endowment process and find evidence of stochastic volatility. Introducing stochastic volatility helps the model generate a more plausible unconditional market price of risk and increases the measured welfare costs of business cycles by about 15%. These asset pricing and welfare results are the result of an agent valuing consumption streams as if a worst case model has generated the data. In our stochastic volatility set up, the robust agent's worst case consumption growth process contains both disasters and a long run risk component.

1 Introduction

One of the most enduring puzzles in the macro-finance literature is the equity premium puzzle. A manifestation of this puzzle is the difficulty of designing a model that simultaneously generates a substantial market price of risk and a low risk free rate, while also respecting stylized facts regarding consumption dynamics, as discussed by Hansen and Singleton (1982) and Mehra and Prescott (1985).

Many attempts to explain the equity premium puzzle have focused on the dynamics of the consumption process. In the long run risks literature, started by Bansal and Yaron (2004), a small but persistent component in consumption growth is used to reconcile the smoothness of consumption with observed risk premia. By definition, direct evidence of the persistent component is hard to detect in post war consumption data, leading to the question of whether or not the component actually exists. An alternate explanation of the equity premium, again based on amending the consumption process, is the idea that agents seek additional premia to compensate for occasional disasters in consumption growth. This approach was first championed by Rietz (1988) and, while sometimes criticized for lack of evidence in US data, has also been promoted in more recent work by Barro (2006) and Barro and Ursua (2008a), Barro and Ursua (2008b). In this paper, we take a different approach. Rather than positing the existence of difficult to detect long run risk or disasters in consumption, we show that the interaction of stochastic volatility in consumption with a fear of model misspecification can generate these phenomena endogenously in the mind of the agent.

The agent in our heteroskedastic endowment economy does not fully trust the joint conditional distribution of the volatility process and consumption series. The agent acknowledges that the model is an approximation to the true data generating process and fears it is misspecified in some way. These fears are expressed through alternative models, or probability distributions, that are distorted versions of the distribution implied by the approximating model and which the agent thinks may be generating the consumption and consumption volatility series. The agent's preference for robustness can be expressed using a two player game between the agent and his alter-ego, or a metaphorical 'evil agent', who attempts to minimize his utility by choosing a worst case, probability distribution. The strength of the agent's preference for robustness is captured by a parameter θ which determines how strongly the evil agent is penalized for choosing distorted distributions that are different from the approximating model, where the degree to which they differ is captured by relative entropy. Thus, θ is used to control how plausible are the misspecifications that the agent implicitly fears - if a distorted process is drastically different from the approximating model then the agent should have already detected this difference and dismissed the process from the set of models that he seeks robustness against. In our model, a changing level of risk in the economy, captured by stochastic volatility in log consumption growth, interacts with an agent's fear that his model is misspecified. This interaction reflects the fact that the level of volatility in the economy affects the agent's ability to detect differences between his approximating model and alternative models. Consequently, the level of volatility affects the set of plausible models the agent entertains as a way of expressing his doubts over his approximating model.

We characterize how the agent's conditional worst case distribution changes depending on the current level of volatility. Specifically, as volatility rises, the marginal entropy penalty the evil agent faces for a given distortion decreases. This leads to a greater distortion of the joint distribution of the innovations to consumption growth and the volatility process in periods with high volatility. The mean of the marginal distribution of the consumption growth innovation varies negatively with the level of volatility. Since the volatility process is persistent, this state dependent mean shift of the consumption growth innovation generates persistence in the worst case consumption growth process. Thus, the agent fears that his consumption growth process has a long run risk component. Similarly, the worst case marginal distribution of the innovation to the stochastic volatility process is also state dependent, with a mean which varies positively with the volatility state. Furthermore, the evil agent chooses to negatively correlate both innovations, making the bad times even worse on average. Unconditionally, because volatility innovations are negatively correlated with consumption innovations, negative skewness is induced in the worst case consumption growth series. Thus this process features disasters more frequently than under the approximating model.

In order to obtain these characterizations of the worst case distributions it is necessary to extend the methodology of robust control to handle general nonlinear and non-Gaussian settings. Our key insight is that if we can compute an approximation to a value function then we can compute an approximation to the worst case model's pdf. The ability to compute an approximation to the pdf is sufficient to be able to draw from this distorted distribution.[1] The ability to draw from this distribution without an *a priori* characterization opens the door for economists to carry out similar analysis in a wide variety of settings where an agent's fear of model misspecification may be important.[2]

[1] The Monte Carlo algorithms described in this paper asymptotically draw from the true distribution, that is, as the number of samples goes to infinity. For any finite sample, they are only approximately distributed from the correct distribution. From here on we will drop the word approximately.

[2] Two related papers Kleshchelski and Vincent (2008) and Drechsler (2010) analyze the effects of stochastic

We apply our methodology in the context of a heteroskedastic endowment process estimated from US consumption data, using Bayesian methods. We present strong evidence of time variation in the conditional standard deviation of log consumption growth and examine its implications for the market price of risk.

Allowing for stochastic volatility enhances the ability of the model to simultaneously generate a high market price of risk and a low risk free rate. That is, for a given value of θ, the model comes closer to attaining the Hansen-Jagannathan bounds than in the homoskedastic case. This is because uncertainty in the estimated parameters induces positive skewness in the distribution of market prices of risk under our posterior. Given the parameter uncertainty reflected in our posterior, we illustrate the proximity of the model's market price of risk to the Hansen and Jagannathan (1991) bounds in terms of a scatter plot of the posterior distribution of the unconditional market price of risk. This depiction yields significant information beyond what can be gleaned from best point or modal estimates and allows us to identify the dimensions of parameter uncertainty that are of particular importance.

Based on our estimates we demonstrate that there is considerable variation in the conditional market price of risk over time and that this variation is primarily due to variation in the conditional variance of the stochastic discount factor, rather than in the conditional mean. We therefore obtain a time varying market price of risk while also generating a risk free rate that is stable over time. From a robustness perspective, we acknowledge that the term 'Market Price of Risk' is something of a misnomer. Under this interpretation, the objects that are typically referred to as the market price of risk and risk premia are not only driven by concerns regarding risk, or gambles over known probabilities, but also fears of model misspecification, where the probabilities over outcomes are unknown. With this in mind, much of the unconditional market price of risk is accounted for by a component that can be interpreted as a market price of model uncertainty and movements over time in the market price of risk reflect, largely, movements in the conditional market price of model uncertainty. We demonstrate that the market price of model uncertainty depends positively on the volatility state of the economy. Given our analysis of the worst case consumption and volatility processes, we interpret the properties of the market price of risk as reflecting the agent's fears of elevated volatility, long run risk and disasters.

Detection error probabilities are used to determine what values of θ are reasonable. De-

volatility in consumption on asset prices in a continuous time setting. Our model is entirely in discrete time. The tools we present can be used in many discrete time representative agent frameworks, a standard workhorse of modern macroeconomics.

tection error probabilities quantify how likely one is to make an error when performing a likelihood ratio test to discriminate between the approximating and worst case models. High detection error probabilities suggest that the competing models are hard to distinguish using the amount of data available and are thus sensible to worry about. We calculate the detection error probabilities for various values of θ and find that the detection error probability curve for the stochastic volatility model lies above that of the homoskedastic model. That is, differences in models are harder to detect in a stochastic volatility setup relative to the homoskedastic model. While the homoskedastic model is able to attain the Hansen-Jagannathan bounds, it does so with a relatively low detection error probability. For values of θ that reach the Hansen-Jagannathan bound, our stochastic volatility model increases the associated detection error probability by a factor of about 3. From an alternative perspective, for a particular detection error probability, the stochastic volatility model allows us to get closer to the Hansen-Jagannathan bounds.

The structure of the paper is as follows. In section 2 we outline the goal of designing a stochastic discount factor that is consistent with the data on returns. In section 3 we describe homoskedastic and heteroscedastic endowment economies that will form the context for our analysis. Section 4 describes our robustness framework and novel techniques to analyze worst case distributions. In section 5 we present our results. In particular we discuss the behavior of the conditional and unconditional market price of risk and show how the worst case model feared by the agent can be used to interpret this behavior. Section 6 concludes.

2 Basic Asset Pricing

A stochastic discount factor is defined as a random variable $\Lambda_{t,t+1}$ that satisfies the pricing relation

$$p_t = E_t[\Lambda_{t,t+1} y_{t+1}], \tag{1}$$

where p_t represents the price at time t of a claim to payoff y_{t+1} in the next period and E_t represents the expectations operator conditional on information at time t. Dividing through by p_t, and defining $R_{t+1} = y_{t+1}/p_t$ we can write this fundamental asset pricing equation as

$$1 = E_t[\Lambda_{t,t+1} R_{t+1}] \tag{2}$$

By taking y_{t+1} to be a conditionally deterministic unit payoff, we observe that the one period risk free rate, R_t^f, is given by the inverse of the expectation of the stochastic discount factor.

Thus we have

$$\frac{1}{R_t^f} = E_t[\Lambda_{t,t+1}] \tag{3}$$

Given an observable sequence of returns, Hansen and Jagannathan (1991) sought bounds on the stochastic properties of the stochastic discount factor. Hansen and Jagannathan showed that an admissible stochastic discount factor must be such that for any zero-price excess return, ξ_{t+1}, the following is true

$$\frac{|E_t[\xi_{t+1}]|}{\sigma_t(\xi_{t+1})} \leq \frac{\sigma_t(\Lambda_{t,t+1})}{E_t[\Lambda_{t,t+1}]}. \tag{4}$$

The quantity on the right hand side of the inequality is typically known as the market price of risk, whereas the left hand side is the Sharpe ratio that captures the additional return on an asset required to compensate for additional undiversifiable risk. We generally work with an unconditional counterpart of this inequality,

$$\frac{|E[\xi_{t+1}]|}{\sigma(\xi_{t+1})} \leq \frac{\sigma(\Lambda_{t,t+1})}{E[\Lambda_{t,t+1}]}.$$

In the absence of a priced risk free asset, the unconditional form of the Hansen-Jagannathan bounds implies a parabola in $(E[\Lambda_{t,t+1}], \sigma(\Lambda_{t,t+1}))$ space such that an assumption on the value of the expectation of the stochastic discount factor pins down the minimal standard deviation for an admissible stochastic discount factor. Clearly, any such assumption on the expectation of the stochastic discount factor is equivalent to an assumption on the value of the risk free rate. We want to design a stochastic discount factor, or a theory of $\Lambda_{t,t+1}$ which lies within these bounds. One theory of $\Lambda_{t,t+1}$ is that it is equal to an agent's intertemporal marginal rate of substitution.[3] To complete this theory of $\Lambda_{t,t+1}$, we need a process for consumption, and a specification of preferences.

3 Two Processes for Consumption

3.1 Homoskedastic and Stochastic Volatility

We first posit that log of consumption follows a random walk with drift ϕ and innovation standard deviation, σ,

$$\begin{array}{lcl} \Delta \log(C_{t+1}) & = & \phi + \sigma\epsilon_{t+1} \\ \epsilon_{t+1} & \sim & N(0,1). \end{array} \tag{5}$$

[3] see Lucas (1978)

This specification has been analyzed in both Tallarini (2000) and Barillas, Hansen, and Sargent (2009) (henceforth BHS). However, there is evidence of time variation in the conditional standard deviation of many macroeconomic series, as documented in Stock and Watson (2002), McConnell and Perez-Quiros (2000), Fernández-Villaverde and Rubio-Ramírez (2007), Justiniano and Primiceri (2008), Bloom, Floetotto, and Jaimovich (2009), Clark (2009), and Ursua (2010) to name but a few. Given the aforementioned evidence, we propose an alternate endowment process that features stochastic volatility in log consumption growth as follows,

$$\begin{array}{lcl} \Delta\log(C_{t+1}) & = & \phi + \sigma\exp(v_{t+1})\epsilon_{1,t+1} \\ v_{t+1} & = & \lambda v_t + \tau\epsilon_{2,t+1} \\ (\epsilon_{1,t+1}, \epsilon_{2,t+1})' & \sim & N(0, I). \end{array} \tag{6}$$

Here, the innovations to the process in $t+1$ have conditional standard deviations that vary over time, driven by the component, v_{t+1}.[4] The process controlling the conditional volatility, v_t is an AR(1) with persistence parameter, λ, and innovations with standard deviation, τ. We use as our consumption series real per capita nondurables and services from 1948:Q2 to 2009:Q4.[5]

3.2 Posteriors from Parameter Estimation

We use Bayesian methods to estimate the parameters of the endowment and volatility process and report them in table 1. We put Uniform[0,1] priors on ϕ, σ, and τ and Uniform[-1,1] on λ and draw from the posterior distribution using a Random Walk Metropolis Hastings algorithm. In our stochastic volatility model, the likelihood is unavailable in closed form, and we therefore rely on the Particle Marginal Metropolis Hastings Algorithm of Andrieu, Doucet, and Holenstein (2010) which entails using a particle filter to evaluate the likelihood within a Metropolis-Hastings Algorithm.[6]

[4] There is a slight technical issue with the data generating process, namely, the tails of the normal distribution are too fat for the level consumption growth process to have finite expectation. This can be corrected by truncating the tails of the normal distribution. The algorithm used to estimate the model would not be affected. Further our perturbation approach described below, entails an n^{th} order approximation which depends on the first n moments of the distribution of the error terms but does not depend on the tails of the distributions per se. Such a distribution could be constructed to have zero mean, and zero third moment, with a variance near 1, and yield nearly identical results.

[5] To compute our log consumption growth data, we added PCESVC96 and PCNDGC96 and and divided by CNP16OV, all obtained from the St. Louis Fed and took log differences.

[6] The PMMH routine converges in distribution for any number of particles. Using 1000 particles for the evaluation of the particle filter. See Smith (2010) for further details regarding this algorithm and specification or Creal (2010) for an overview of sequential monte carlo methods in economics and finance. We use 200000 draws from the posterior to compute the above tables. In the following sections, we only rely on samples of the posterior distribution of parameters and do not use all 200000 draws.

Table 1: Posterior Distribution of Parameters

	Homoskedastic		Stochastic Volatility	
	Mean	(0.025,0.975)	Mean	(0.025,0.975)
ϕ	0.0047	(0.0040,0.0053)	0.0047	(0.0041,0.0053)
σ	0.0052	(0.0048,0.0057)	0.0047	(0.0036,0.0066)
λ			0.8624	(0.6770,0.9701)
τ			0.2109	(0.1165,0.3318)

The posteriors for the homoskedastic model are centered around similar values to those found in previous studies, with a slightly lower growth rate and higher variance, most likely due to the longer data series including the recession starting in 2007 Q4. Our stochastic volatility model yields a similar growth rate of consumption as the homoskedastic model. We find that the σ parameter, although not directly analogous across models, is estimated far more imprecisely than in the homoskedastic case. The volatility sequence is persistent with an autoregressive parameter λ around 0.86, and has a conditional standard deviation of about 0.21. In both models there is a reasonable degree of parameter uncertainty, reflected in the dispersion of our posteriors. Clearly this dispersion will imply uncertainty in our beliefs regarding certain model properties and moments. In fact, as we shall discuss below, it is instructive to see what dimensions of parameter uncertainty are important for particular model moments.

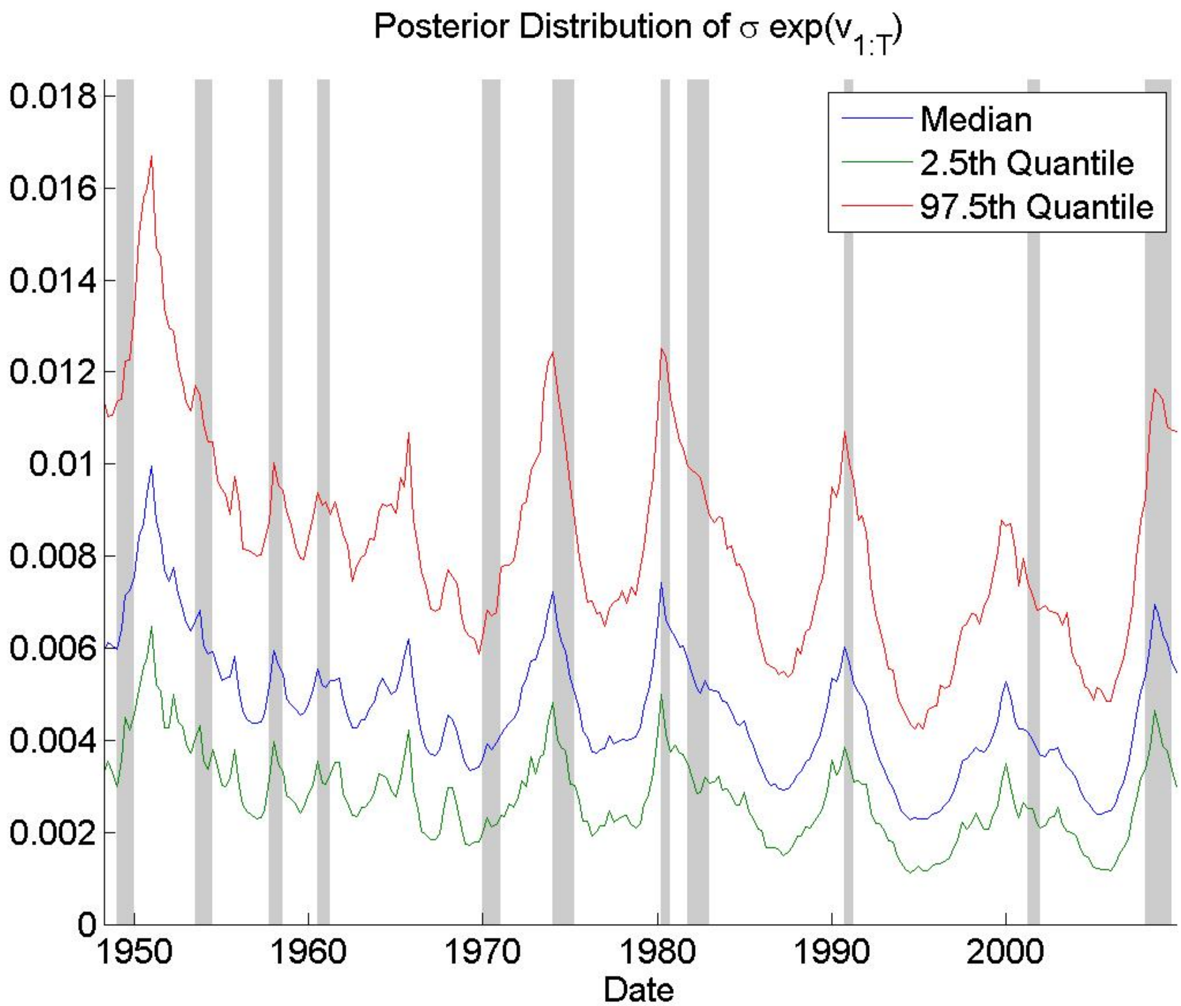
We plot the posterior distribution of volatility in figure 1. We notice that there is indeed much movement in this volatility sequence. A general decline is observed throughout the sample period and although seemingly acyclical for the earlier sample the series appears to peak at the beginning of recessions experienced after the mid-1970s. In addition, the results of likelihood ratio tests and calculation of Bayes factors suggest strongly that stochastic volatility is present.[7]

4 A Robust Specification of Preferences

4.1 Martingales, Martingale Increments and Distorted Distributions

Within the robust control literature, it is convenient to have a language for characterizing alternative models in relation to a maintained approximating model. A robust agent is

[7] The log likelihoods are 963.0 for the model with stochastic volatility and 947.9 for the homoskedastic model. Estimates of the log Bayes factor ranged from 10.9824 to 11.0812, yielding further evidence in favor of the stochastic volatility model.



Posterior Distribution of σ exp($v_{1:T}$)
0.018
0.016
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
1950
1960
1970
1980
1990
2000
Date
Median
2.5th Quantile
97.5th Quantile

Figure 1:

concerned with how his welfare behaves across a set of plausible models. The fact that he considers a set of models reflects his doubts regarding his approximating model, which he suspects is misspecified in some way. Regarding the equilibrium of an economic model as a probability distribution it is natural to specify alternative models in terms of distortions of the distributions associated with the agent's approximating model. In Hansen and Sargent (2008), it is proposed that the distortions be characterized in terms of Martingales. These Martingales act as Radon-Nikodym derivatives by twisting the measures implicit in the approximating model so as to obtain absolutely continuous measures that represent alternative models considered by the agent.

Formally, let $\Im_t$ be information available at t and define a non-negative $\Im_t$ measurable function M_t such that $E[M_t|\Im_0] = 1$. This function can be used to derive a probability measure that is absolutely continuous with respect to the measure over $\Im_t$ implied by the approximating model. With respect to the undistorted measure, M_t is a Martingale. Using this Martingale we can define distorted expectations as follows,

$$\tilde{E}[W_t] \doteq E[M_t W_t]. \tag{7}$$

As a measure of how different the distorted measure is from the undistorted measure associated with the approximating model, we use the concept of relative entropy, conditional on time-zero information $E[M_t log(M_t)|\Im_0]$.

Analytically and, as we shall see, computationally, it is convenient to factorize the Martingale, M_t, into a sequence of increments, m_t as, $m_{t+1} = \frac{M_{t+1}}{M_t}$ if $M_t > 0$ and 1 otherwise. Thus $M_{t+1} = m_{t+1} M_t$, and $M_t = M_0 \prod_{j=1}^{T} m_t$. Using this Martingale increment we can define a distorted conditional expectation for a $\Im_{t+1}$-measurable random variable, b_{t+1}, given $\Im_t$ and, more generally, use m_{t+1} to capture the distortion of the conditional distribution of $\Im_{t+1}$ given $\Im_t$

$$\frac{E[M_{t+1} b_{t+1}|\Im_t]}{E[M_{t+1}|\Im_t]} = \frac{E[M_{t+1} b_{t+1}|\Im_t]}{M_t} = E[m_{t+1} b_{t+1}|\Im_t]. \tag{8}$$

4.2 Multiplier Preferences

We adopt Hansen-Sargent multiplier preferences as a way of expressing the agent's desire for robustness to model misspecification. Let x_t denote the state at time t, ϵ_t be a random vector with density $p(\epsilon_t)$. Given an initial value of the state x_0 we can compute the state as $x_t = f_t(\epsilon_{1:t}, x_0)$. The value function under these preferences, in our endowment case, is given by

$$W(x_0) = \min_{\{m_{t+1}\}} \sum_{t=0}^{\infty} E\{\beta^t M_t \left[log(C(x_t)) + \beta\theta E(m_{t+1} log(m_{t+1})|\epsilon_{1:t}, x_0)\right] |x_0\}.$$

Where the minimization is subject to the evolution of the endowment process and $M_{t+1} = m_{t+1}M_t$, $E[m_{t+1}|\epsilon_{1:t}, x_0] = 1$, $m_{t+1} \geq 0$ and $M_0 = 1$. The agent's desire for robustness is reflected in the minimization over the sequence of Martingale increments. The degree of robustness is controlled by the penalty parameter θ that enters in the objective by multiplying the conditional relative entropy associated with a given distortion. For $\theta > 0$, the agent is penalized for considering distortions from his approximating model. Thus, a particularly implausible distorted model may imply dynamics that are painful for the agent but its negative effect on the objective is offset by a positive countervailing contribution reflecting its high entropy and, thus, it does not solve the minimization problem.

We seek a recursive expression of the problem. First, let the current state be x, and the state tomorrow be $x(\epsilon)$. The Bellman equation for the problem defined above is then

$$W(x) = log(C(x)) + \min_{m(\epsilon;x)\geq 0} (\beta \int [m(\epsilon;x)W(x(\epsilon)) + \theta m(\epsilon;x)log(m(\epsilon;x))]p(\epsilon)d\epsilon). \quad (9)$$

where $p(\epsilon)$ is the density of the random variable ϵ under the approximating model and the minimization is subject to $\int m(\epsilon;x)p(\epsilon)d\epsilon = 1$. We have written the minimizing likelihood ratio $m_{t+1}(\epsilon_{t+1})$ as a time invariant function of the state, $m(\epsilon;x)$ to allow this worst case distribution to be state dependent. If we substitute the solution to the minimization problem into the Bellman equation we obtain the expression

$$W(x) = log(C(x)) - \beta\theta log\left(\int exp\left(\frac{-W(x(\epsilon))}{\theta}\right)p(\epsilon)d\epsilon\right). \quad (10)$$

This indirect utility function of an agent with multiplier preferences is algebraically identical to that of an agent with risk sensitive preferences who fully trusts the approximating model. One can map between a penalty parameter θ, and a risk aversion parameter γ under the risk sensitive preferences interpretation.[8] As discussed by BHS, the difference lies in the interpretation of the θ. In the risk sensitive case the parameter reflects attitudes towards well defined, quantifiable risk whereas under the robustness interpretation it reflects the degree to which the agent fears model misspecification.

We can view this specification of robust preferences as a particular form of a two player game, in which the agent discussed so far attempts to maximize his welfare, but is partly thwarted by an 'evil agent' who distorts the model facing the agent. The extent to which the evil agent can distort the model is constrained by restricting the distorted models to be similar to the undistorted model. It is useful to employ this language in our context where the

[8]The mapping in this case is $\theta = \frac{-1}{(1-\beta)(1-\gamma)}$.

evil agent's actions can be thought of as the choice of the distorting Martingale increment, similarity is defined in terms of entropy and the extent to which the evil agent is restricted is captured by θ.

The minimizing Martingale increment for the above problem, which can be thought of as the evil agent's policy function, is given by

$$\hat{m}_{t+1}(\epsilon_{t+1}) = \left(\frac{exp\left(\frac{-W(x(\epsilon_{t+1}))}{\theta}\right)}{E_t\left[exp\left(\frac{-W(x(\epsilon_{t+1}))}{\theta}\right)\right]} \right). \tag{11}$$

An agent with multiplier preferences has a stochastic discount factor of the following form,

$$\Lambda_{t,t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-1} \left(\frac{exp\left(\frac{-W_{t+1}}{\theta}\right)}{E_t\left[exp\left(\frac{-W_{t+1}}{\theta}\right)\right]} \right). \tag{12}$$

Provided that $\theta < \infty$ we observe that the stochastic discount factor comprises two components,

$$\Lambda_{t,t+1} = \Lambda^R_{t,t+1}\Lambda^U_{t,t+1},$$

where we define,

$$\begin{aligned} \Lambda^R_{t,t+1} &= \beta \left(\frac{C_{t+1}}{C_t}\right)^{-1} \\ \Lambda^U_{t,t+1} &= \frac{exp\left(\frac{-W_{t+1}}{\theta}\right)}{E_t\left[exp\left(\frac{-W_{t+1}}{\theta}\right)\right]}. \end{aligned}$$

The first component, $\Lambda^R_{t,t+1}$, takes the form of the stochastic discount factor derived from time separable logarithmic preferences. The second component, $\Lambda^U_{t,t+1} (= \hat{m}_{t+1}(\epsilon_{t+1}))$ is a wedge in the fundamental asset pricing equation reflecting the distortion of the approximating model's conditional distribution such that payoffs are evaluated under the worst of a set of plausible models considered by the agent. The robust agent uses his stochastic discount factor to price assets as,

$$\begin{aligned} E_t\left[\Lambda_{t,t+1}R_{t+1}\right] &= 1 \\ \beta \int R(\epsilon) \left(\frac{C(x(\epsilon))}{C(x)}\right)^{-1} \frac{exp\left(\frac{-W(x(\epsilon))}{\theta}\right)}{E_t\left[exp\left(\frac{-W(x(\epsilon))}{\theta}\right)\right]} p(\epsilon)d\epsilon &= 1 \\ \beta \int R(\epsilon) \left(\frac{C(x(\epsilon))}{C(x)}\right)^{-1} \tilde{p}(\epsilon)d\epsilon &= 1 \\ \tilde{E}_t\left[\beta \left(\frac{C_{t+1}}{C_t}\right)^{-1} R_{t+1}\right] &= 1. \end{aligned}$$

Thus, an agent who fears his models is misspecified prices assets as if he has logarithmic period utility, but uses a distorted conditional expectations operator $\tilde{E}_t$ and an associated distorted density $\tilde{p}(\epsilon_t)$.

4.2.1 Calibrating θ

While there is an algebraic equivalence between the value function of a risk sensitive agent and the indirect utility function of an agent with multiplier preferences, very different thought experiments govern what parameterization is considered plausible. Under the risk sensitive interpretation, the choice of γ (risk aversion) can be disciplined by experimental evidence and introspection (*à la* Arrow-Pratt). Under the robustness interpretation, we discipline θ by detection error probabilities. Detection error probabilities characterize a set of distorted models in terms of whether or not, with a limited amount of data, an agent could accurately distinguish between the worst case and approximating models using likelihood ratio tests. Using a homoskedastic model, BHS show that large values of risk aversion, γ that attain the Hansen-Jagannathan bound can be mapped to a penalty parameter, θ, that is associated with a detection error probability between 1 and 5%.

4.3 Robust Control in General Environments

To proceed, we need to solve the problem of how to implement robust control analysis in our environment. Previous results in the robust control literature relied on analytical solutions or, in the Linear-Quadratic case, Ricatti Equation algorithms, to compute explicit characterizations of value functions, likelihood functions, and the worst case model $\tilde{p}(\epsilon_t)$. Our nonlinear environment does not allow us to compute explicit characterization of these model objects directly. Nevertheless, we propose to work with an approximation to the value function and associated objects and take a novel approach to characterizing the worst case model and likelihood function.

We choose to obtain a higher order perturbation approximation to the value function although one could theoretically use value function iteration or projection methods. Below we show how to evaluate a likelihood function and draw from the worst case model, even though cannot directly compute an explicit characterization of the worst case. If one were able to write out the functional form of $\tilde{p}(\epsilon_t)$, it would be unrecognizable as any known pdf. However, because we can evaluate (an approximation to) this object, we can draw from it using Monte Carlo methods. We need to be able to perform these tasks in order to compute detection error probabilities and to evaluate the set of models that our agent seeks robustness against. We first describe in general terms our procedure and then discuss the details in the following section.

Summary of Method to draw from Worst Case model

1. Partition the state, x, into random elements, ϵ, and a predetermined variables s, such that $s' = f(x)$
2. Compute an approximation to the value function, using value function iteration, perturbation, projection, etc
3. Compute $W(x') = W(\epsilon', f(x))$
4. Compute $m(x') \propto exp(\frac{-W(x')}{\theta})$
5. Compute $\tilde{p}(\epsilon'; x) \propto m(x')p(\epsilon')$
6. Use a Monte Carlo algorithm to draw (approximately) from $\tilde{p}(\epsilon'; x)$

4.3.1 Drawing from the Worst Case Distribution

Here we describe a strategy to characterize this worst case model in general settings. Let us partition the state vector into components that are unknown until the resolution of randomness in t, ϵ_t, and those that are predetermined in t, s_t. Thus, the state, $x_t = (\epsilon_t, s_t)$. In the case of our approximate solution of the stochastic volatility model, it is useful to think of the predetermined component being the lagged volatility level, $s_t = \{v_{t-1}\}$.[9] The stochastic component is the innovation in the endowment process and the innovation in the stochastic volatility process, $\epsilon_t = \{\epsilon_{1,t}, \epsilon_{2,t}\}$.

In general, one does not have a closed form solution to the value function $W(x)$. We approximate the value function by a third order Taylor polynomial around the deterministic steady state. Perturbation is much faster than value function iteration or Chebyshev polynomials, though not as accurate.[10] In practice, almost any solution method would work. Using this approximation to the value function, we can then compute an approximation to the minimizing likelihood ratio chosen by the evil agent.

We denote the perturbation-based approximation of the minimizing Martingale increment $\tilde{m}(x_t)$ with our partitioned state vector as $x_t = (\epsilon_t, s_t)$ where ϵ_t are the shocks realized in t and s_t are the components of the state that are predetermined in t. The transition density is

$$p(X_t|X_{t-1}) = p(\epsilon_t),$$

[9]Our actual solution includes lagged state variables such as the lagged continuation value, which is needed for accounting purposes for the functional objects we examine in the perturbation solution. The variables in s_t should include anything that evolves deterministically given the current state, ie $s_{t+1} = f(\epsilon_t, s_t)$.

[10]See Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2006)

since the predetermined components of the state evolve as $s_t = f(\tilde{\epsilon}_{t-1}, s_{t-1})$. Similarly the distorted transition density is

$$\tilde{p}(X_t|X_{t-1}) = \tilde{p}(\epsilon_t; s_t) = \tilde{p}(\epsilon_t; f(s_{t-1})).$$

We can express the distorted distribution of ϵ_t as

$$\tilde{p}(\epsilon_t; s_t) = \tilde{m}(\epsilon_t; s_t)p(\epsilon_t), \tag{13}$$

where we recall that $p(\cdot)$ is the density of the approximating model and $\tilde{p}(\cdot)$ represents the twisted probability density function associated with the worst case distribution. This decomposition shows clearly how the distorted distribution can depend on the predetermined states, and thus may vary though time. Also note that in our discrete time nonlinear model, the distorted distribution may imply that the shocks are dependent, despite the fact that under the undistorted model, the shocks are independent.

Our insight is that even though we do not know anything about this worst case distribution, since we are able to evaluate the distorted pdf, we can draw from it using Monte Carlo methods. Using the draws from the distorted distribution, we can then compute any relevant quantity that deals with this worst case distribution. Here we propose a few simple algorithms for drawing from the distorted distribution. Being able to draw from the worst case model allows us to approximate $\tilde{p}(dx_t) = \sum_{i=1}^{N} \delta_{\tilde{x}_t^i}(dx_t)$ and any distorted moment of interest, as well as generate a sequence of data under the worst case, $\tilde{Y}_{1:T}$.

4.3.2 Random Walk Metropolis - Hastings

Our first method of drawing from the distorted distribution is a simple random walk Metropolis-Hastings algorithm.[11] This algorithm produces a set of correlated draws $\tilde{\epsilon}_t \sim \tilde{p}(\epsilon_t, s_t)$ according to a Markov Chain whose invariant distribution is the desired target distribution. In order to produce a sample of length T, we pick one draw and repeat for each time period. The algorithm is as follows: Given we have some $\tilde{\epsilon}_{t-1}, s_{t-1}$

1. Set $s_t = f(\tilde{\epsilon}_{t-1}, s_{t-1})$
2. for $i = 1, \ldots N$
    - Draw $\epsilon_t^* \sim q(\epsilon_t^{i-1}, \cdot)$
    - Set $\epsilon_t^i = \epsilon_t^*$ with probability $min\left\{1, \frac{\tilde{p}(\epsilon_t^*)q(\epsilon_t^*, \epsilon_t^{i-1})}{\tilde{p}(\epsilon_t^{i-1})q(\epsilon_t^{i-1}, \epsilon_t^*)}\right\}$ and $\epsilon^i = \epsilon_t^{i-1}$ otherwise

[11] See Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953) or Hastings (1970).

3. draw one ϵ_t from $\{\epsilon_t^i\}_{i=1}^N$

4. increase t

This algorithm at each time period produces a Markov Chain whose invariant distribution is the distribution we want to sample from. These draws can then be used to estimate state by state moments, or, by repeating for $t = 1, \ldots, T$ this produces a series of states $(\epsilon_{1:T}, s_{1:T})$ which can be used to construct data under the distorted distribution.

4.3.3 Sampling Importance Resampling

Our second method is based on the Sampling Importance Resampling (SIR) algorithm of Rubin (1987) and Smith and Gelfand (1992) for drawing from this worst case model. We simulate draws from the approximating model, compute the associated importance weights (which in this case are given by the minimizing martingale increment) and then resample with replacement according to those weights. This yields an unbiased estimate of the distribution $\tilde{p}(\epsilon_t, s_t)$. Again, by repeating these steps one can generate a series of states from the distorted distribution and construct data under the worst case model. Given some $\tilde{\epsilon}_{t-1}, s_{t-1}$

1. set $s_t = f(\tilde{\epsilon}_{t-1}, s_{t-1})$

2. draw $\epsilon_t^i \sim p(\epsilon_t)$, $i = 1, \ldots, N$

3. assign weight $W_t^i = \tilde{m}(\epsilon_t; s_t) = exp(\frac{-W(x_t)}{\theta})$

4. resample with probability $\propto W_t^i$, call new set of draws $\tilde{\epsilon}_t$

5. draw once from $\{\epsilon_t^i, s_t\}_{i=1}^N$ and store

6. increase t

We use this algorithm to compute the moments and generate the distorted consumption series in the sections below. Note that while we use perturbation for the value function and SIR to draw from the worst case model, nearly any combination of value function approximation and Monte Carlo algorithm could be used to produce draws from the worst case model.

4.3.4 Evaluating Likelihood Functions

We have a nonlinear state space system with a likelihood function that is unavailable in closed form. Here we describe how we compute likelihood functions and ultimately detection error probabilities and extend the techniques of Anderson, Hansen, and Sargent (2003) to nonlinear discrete time models. Although our innovations are normally distributed under the approximating model, the strategy outlined below can handle any probability distribution. Further, the distorted model in a nonlinear framework may no longer be Gaussian even if the approximating model is. Let $Y_{1:t}$ be a vector of observables, or signals, up to time t and $X_{1:t}$ be a vector of unobservables up to time t. We, as an agent or econometrician, want to calculate,

$$\begin{aligned} p(Y_{1:T}|\theta) &= \textstyle\prod_{t=1}^{T} p(Y_t|Y_{1:t-1}) \\ &= \textstyle\prod_{t=1}^{T} \int p(Y_t|X_t)p(X_t|Y_{1:t-1})dX_t. \end{aligned}$$

We can use a particle filter to do so. For more uses of particle filtering in economics, see Fernández-Villaverde and Rubio-Ramírez (2007), Flury and Shephard (2008), Creal (2010).

Given a set of draws $\{X^i_{1:t-1}\}_{i=1}^N$ with weights $\{W^i_{t-1}\}_{i=1}^N$

- draw $X^i_t \sim q(X^i_t|Y_t, X^i_{1:t-1})$
- compute weight $W^i_t = \frac{p(Y_t|X^i_t)p(X^i_t|X^i_{t-1})}{q(X^i_t|Y_t,X^i_{1:t-1})} W^i_{t-1}$
- compute likelihood contribution as $L_t = \sum_{i=1}^N W^i_t$
- resample $X^i_{1:t} = (X^i_{1:t-1}, X^i_t)$ with probability $\propto W^i_t$, set weights $W^i_t = 1/N$

The likelihood under the approximating model is then $p(Y_{1:T}|\theta) = \prod_{t=1}^T L_t$. When importance density, $q(X^i_t|Y_t, X^i_{1:t-1})$ is equal to the transition density, $q(X^i_t|Y_t, X^i_{1:t-1}) = p(X_t|X_{t-1})$, the weights are $\propto p(Y_t|X^i_t)$. We need to make the following strong assumption for tractability, that the evil agent cannot distort signals to the econometrician. That is,

$$\tilde{p}(Y_t|X_t) = p(Y_t|X_t).$$

In order to make the detection error probabilities operational for our stochastic volatility model, we need to add measurement error to the consumption sequence as follows

$$\Delta log(C_t) = \phi + \sigma e^{v_t}\epsilon_{1,t} + \sigma_{me}\epsilon_{3,t}.$$

It is often assumed that the shocks to the state equation (in our case $\epsilon_{2,t}$) are independent of shocks to the measurement equation ($\epsilon_{1,t}$) and that one has the ability to compute the probability density functions of the associated randomness. Under our distorted probability distribution, this is no longer the case as we have allowed the evil agent to distort the joint distribution of $(\epsilon_{1,t}, \epsilon_{2,t})$, and we have no way to calculate the marginal density functions or re-write these shocks in such a way that they are independent of each other. Thus, by adding measurement error to the state equation, we can redefine what is the state in our stochastic volatility model, $X_t = (\epsilon_{1,t}, \epsilon_{2,t}, v_{t-1})$, and then rely on standard techniques in the nonlinear filtering literature. Using measurement errors as a tool to evaluate likelihoods has a long history in economics, especially with estimating production economies.[12] We constrain σ_{me} to be an order of magnitude smaller than σ, so that $\sigma_{me} = 0.1\sigma$. Now we have the ability to evaluate the likelihood under the distorted probability measure. Here we wish to calculate

$$\begin{aligned}
\tilde{p}(Y_{1:T}|\theta) &= \prod_{t=1}^{T} \tilde{p}(Y_t|Y_{1:t-1}) \\
&= \prod_{t=1}^{T} \int \tilde{p}(Y_t|X_t)\tilde{p}(X_t|Y_{1:t-1})dX_t \\
&= \prod_{t=1}^{T} \int p(Y_t|X_t)\tilde{p}(X_t|Y_{1:t-1})dX_t.
\end{aligned}$$

The key idea is to use the approximating model as the proposal distribution in the particle filtering step, as follows:

Given a set of draws $\{X^i_{1:t-1}\}_{i=1}^N$ with weights $\{W^i_{t-1}\}_{i=1}^N$

- draw $X^i_t \sim p(X^i_t|X^i_{t-1})$
- compute weight
$$\begin{aligned}
W^i_t &= \frac{p(Y_t|X^i_t)\tilde{p}(X^i_t|X^i_{t-1})}{p(X^i_t|X^i_{t-1})} W^i_{t-1} \\
&= p(Y_t|X^i_t)\tilde{m}(X^i_t)W^i_{t-1}
\end{aligned}$$
- compute likelihood contribution as $\tilde{L}_t = \sum_{i=1}^N W^i_t$
- resample $X^i_{1:t} = (X^i_{1:t-1}, X^i_t)$ with probability $\propto W^i_t$ and set $W^i_t = 1/N$

The distorted likelihood is then $\tilde{p}(Y_{1:T}|\theta) = \prod_{t=1}^{T} \tilde{L}_t$. The sampling weights are multiplied by the martingale distortion term $\tilde{m}(X^i_t)$ relative to the case where we could draw from the distorted transition density $\tilde{p}(X_t|X_{t-1})$. We do not need to use the approximating model, but it seems a natural proposal distribution to choose. If a different importance density is used, the weights become,

$$W^i_t = \frac{p(Y_t|X^i_t)\tilde{m}(X^i_t)p(X^i_t|X^i_{t-1})}{q(X^i_t|Y_t, X^i_{1:t-1})} W^i_{t-1}$$

but the rest of the algorithm remains the same.

[12] We do not try to re-estimate the model including measurement error.

5 Results

Given our posteriors for the parameters of the endowment processes we are able to calculate posteriors for various moments of interest. In the case of the homoskedastic model this typically entails substituting draws from our posterior into closed form expressions. In the case of the stochastic volatility model, where closed from expressions for moments are generally unavailable, we will simulate the economy under each parameter draw to obtain simulation-based posteriors for moments. In the latter case, we therefore confront errors arising both from the use of perturbation approximation and from sampling variability. We choose to represent our results in terms of a parameter γ and label those multiplier preferences MP (recall $\theta = \frac{-1}{(1-\beta)(1-\gamma)}$). For a $\gamma \in \{2, 15, 35, 50\}$, we have an agent who fears his model is misspecified with a penalty parameter $\theta \in \{200, 14.28, 5.88, 4.08\}$ respectively, or as we will see in the stochastic volatility case, a detection error probability of approximately $p(\theta) \in \{0.5, 0.3, 0.10, 0.05\}$.

5.1 Unconditional Market Price of Risk

We begin this section by examining the implications of the homoskedastic model. In figure 2 we plot the average under our posterior of $(E[\Lambda_{t,t+1}], \sigma(\Lambda_{t,t+1}))$ pairs for both the expected utility and multiplier preferences cases without stochastic volatility. Figure 2 is, allowing for the slightly different sample period used, comparable to the graphs of Tallarini discussed and replicated in BHS.

We first note the difficulty of attaining the Hansen-Jagannathan bounds using expected utility and remind the reader that the slope of a ray from the origin going through these points is the market price of risk in the economy. While the mean point estimates of the market price of risk rises under the expected utility, the risk free rate rises as well, thereby preventing the Hansen-Jaganathan bound from being reached. However, with multiplier preferences, raising the value of γ implies a stochastic discount factor with properties that allow the $(E[\Lambda_{t,t+1}], \sigma(\Lambda_{t,t+1}))$ pairs to approach the bound. We also list our posterior means for the the market price of risk in table 2. Note that both expected utility and multiplier preferences imply the same (constant) market price of risk.

In section 2, we found evidence of stochastic volatility in the consumption process. It is therefore natural to re-examine the graphs of Tallarini and BHS while accounting for heteroskedasticity. In figure 3, we compare means of the $(E[\Lambda_{t,t+1}], \sigma(\Lambda_{t,t+1}))$ pairs now computed using draws from the posterior distribution of the stochastic volatility model. For

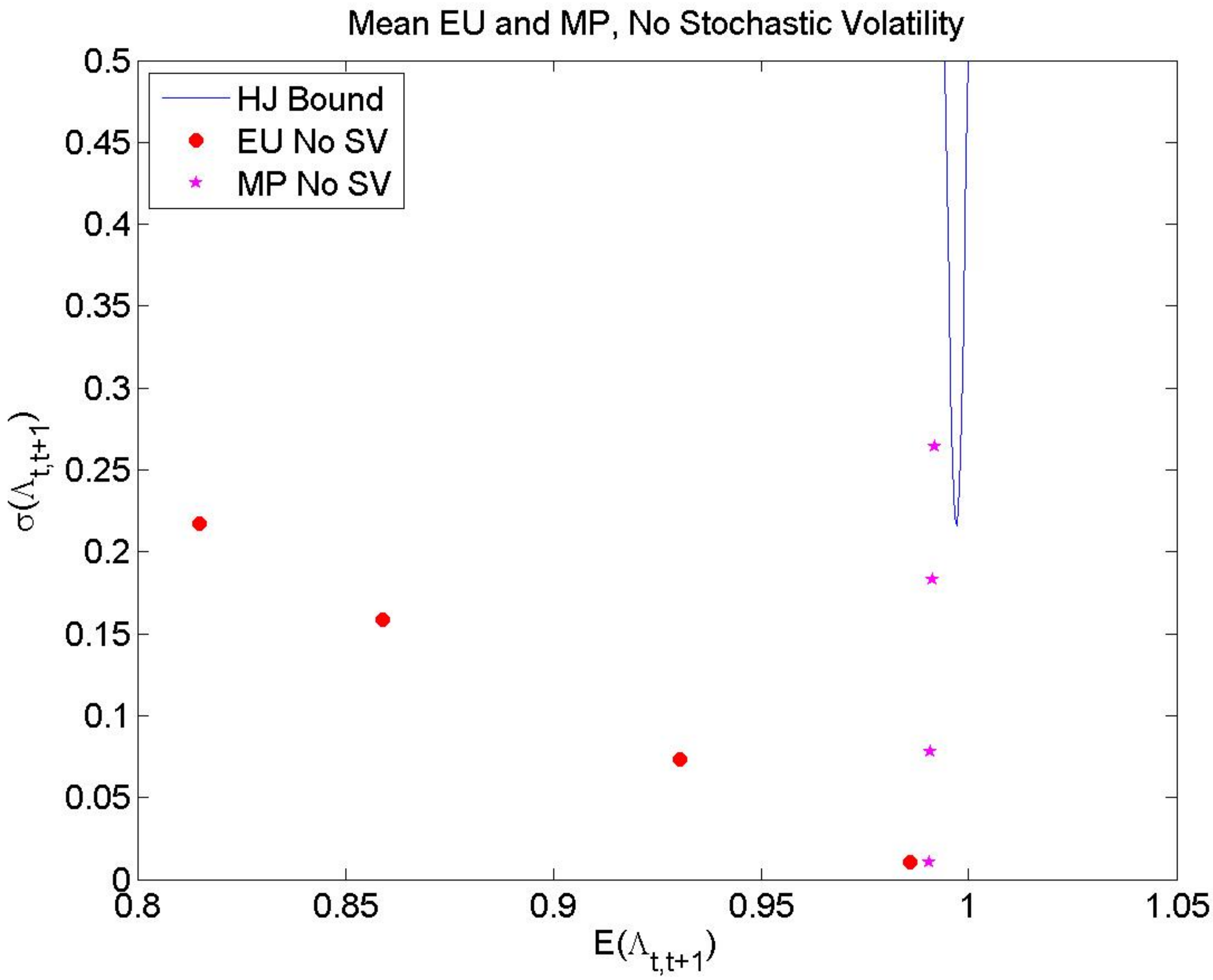


Figure 2: The Graph of Tallarini

Table 2: Multiplier Preferences Without Stochastic Volatility

γ	Mean	(0.025,0.975)
2	0.0105	(0.0096,0.0115)
15	0.0787	(0.0720,0.0861)
35	0.1850	(0.1689,0.2024)
50	0.2667	(0.2431,0.2923)

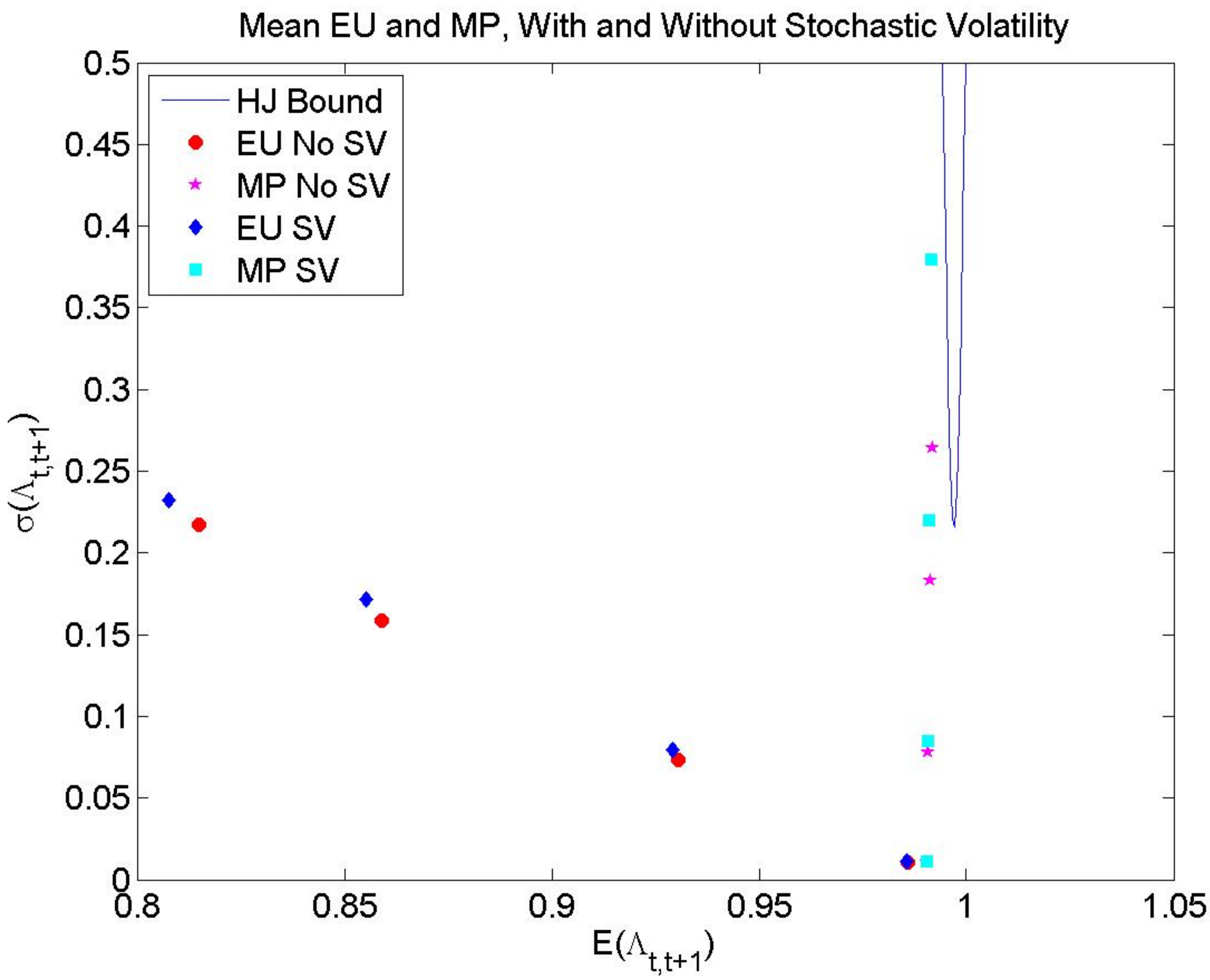


Figure 3: Hansen Jagannathan Bounds With Stochastic Volatility

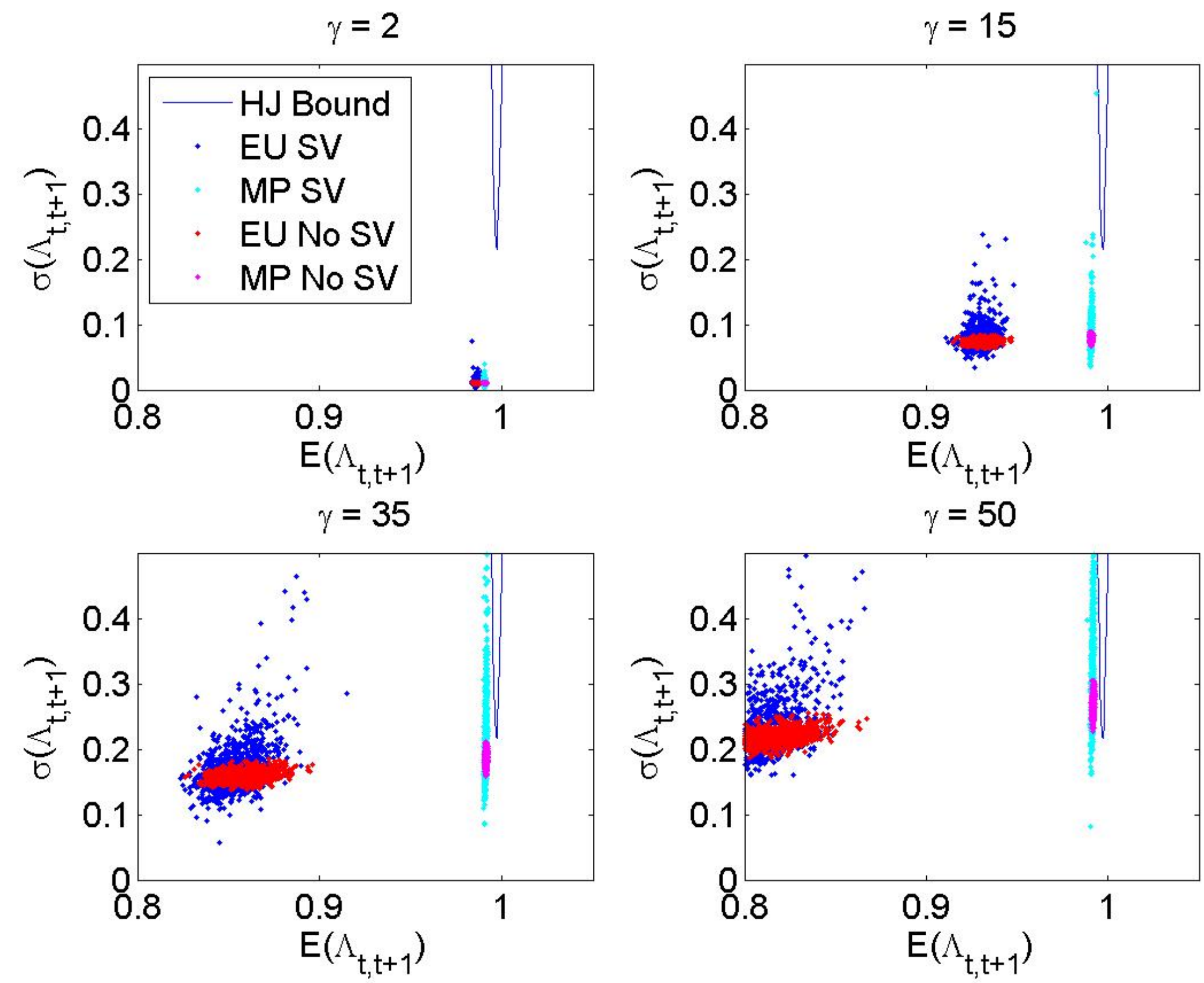


Figure 4: Hansen Jagannathan Clouds With Stochastic Volatility

each parameter draw, due to the lack of closed form expressions for these objects, we solve the model using perturbation and simulate to obtain these moments.[13]

We notice that in the expected utility case, allowing for the possibility of stochastic volatility leads to estimates that imply a slightly higher risk free rate and market price of risk, relative to those implied by the homoscedastic model estimates. In the multiplier

[13] In our perturbation solution we define objects $E_t[\Lambda_{t,t+1}]$, and $\Lambda_{t,t+1}$, simulate the model to compute $E[\Lambda_{t,t+1}] = \frac{1}{T}\sum_{t=1}^{T} E_t[\Lambda_{t,t+1}]$ and $\sigma(\Lambda_{t,t+1}) = \text{sqrt}(\frac{1}{T}\sum_{t=1}^{T}[(\Lambda_{t,t+1} - \frac{1}{T}\sum_{j=1}^{T}\Lambda_{j,j+1})^2])$. Estimating the unconditional mean $E[\Lambda_{t,t+1}]$ using $\frac{1}{T}\sum_{t=1}^{T}\Lambda_{t,t+1}$ yields much more volatile and inaccurate estimates of the mean compared to using the average of the conditional expectation which is why it was used. This is because $\Lambda_{t,t+1}$ is much more volatile than $E_t[\Lambda_{t,t+1}]$. We use $E[\Lambda_{t,t+1}] = \frac{1}{T}\sum_{t=1}^{T}\Lambda_{t,t+1}$ to calculate $\sigma(\Lambda_{t,t+1})$ to center the estimate used in the variance calculation.

Table 3: Market Price of Risk Estimates with Stochastic Volatility

	Multiplier Preferences		Expected Utility	
γ	Mean	(0.025,0.975)	Mean	(0.025,0.975)
2	0.0112	(0.0077,0.0174)	0.0113	(0.0084,0.0170)
15	0.0860	(0.0587,0.1437)	0.0852	(0.0631,0.1294)
35	0.2218	(0.1383,0.4803)	0.2002	(0.1452,0.3118)
50	0.3827	(0.2079,0.8819)	0.2867	(0.2087,0.4561)

preferences case, we observe a large increase in the average market price of risk without any noticeable penalty to the risk free rate.

To explain mechanically why this occurs, we plot $(E[\Lambda_{t,t+1}], \sigma(\Lambda_{t,t+1}))$ pairs for individual parameter draws from our posteriors. This visual representation, seen in figure 4, illustrates more clearly how our parameter uncertainty translates into uncertainty over the market price of risk. We see that the introduction of stochastic volatility results in greater dispersion in both the expected utility and the multiplier preferences cases.[14] The dispersion relative to the homoskedastic case is primarily in the $\sigma(\Lambda_{t,t+1})$ dimension. In particular, the long positive tail in the posterior distribution of market prices of risk under the heteroskedastic estimates leads to the posterior mean of the market price of risk being pulled upwards, especially in the multiplier preferences case. Thus allowing for stochastic volatility in the presence of multiplier preferences helps us approach the Hansen-Jaganathan Bounds for a substantially lower value of γ than in the homoskedastic case. Indeed, comparing tables 2 and 3, for $\gamma = 35$ and $\gamma = 50$ the increases in the mean market price of risk are of approximately 20 and 35 percent, respectively.

5.2 Conditional Asset Pricing

In the homoskedastic endowment case the market price of risk is constant. However, in the presence of stochastic volatility, the market price of risk varies over time. Since we employ a third order perturbation approximation, we are able to capture this time variation. Due to the failure of expected utility to reach the unconditional Hansen-Jagannathan bound, we will focus on multiplier preferences.

In figure 5, we plot the mean of the joint posterior distribution of $E_t[\Lambda_{t,t+1}]$ and $\sigma_t(\Lambda_{t,t+1})$ at different state configurations and with $\gamma = 50$. For each parameter draw, we obtain

[14] Some of the increased dispersion in the clouds reflects errors of approximation and simulation of the model.

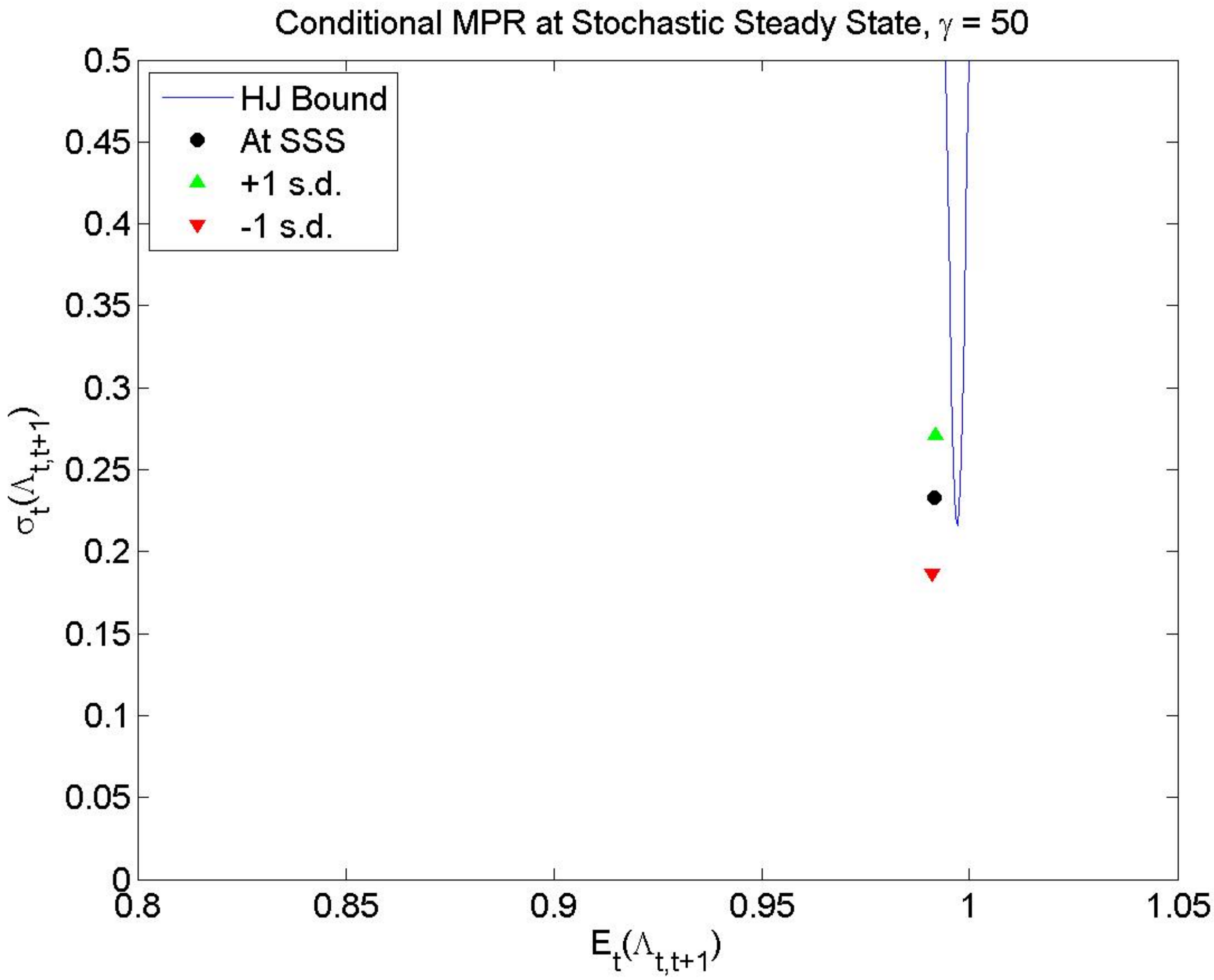


Figure 5: Conditional Market Price of Risk at Steady State: Effect of Volatility Innovations

Table 4: Conditional MPR, effect of $\epsilon_{2,t}$

ϵ_2	Mean	(0.025,0.975)
+1	0.2736	(0.2082,0.3714)
0	0.2352	(0.1783,0.3219)
-1	0.1885	(0.1352,0.2649)

Table 5: Variation in Conditional Moments of $\Lambda_{t,t+1}$, $\gamma = 50$

	Mean	(0.025,0.975)
$\sigma(E_t[\Lambda_{t,t+1}])$	0.00083	(0.00038 , 0.00216)
$\sigma(R_t^f)$	0.00085	(0.00039 , 0.00220)
$\sigma(E_t[\Lambda_{t,t+1}^R])$	0.00001	(0.00000 , 0.00002)
$\sigma(V_t[\Lambda_{t,t+1}^R]$	0.00002	(0.00001 , 0.00004)
$\sigma(E_t[\Lambda_{t,t+1}^U])$	0.00000	(0.00000 , 0.00000)
$\sigma(V_t[\Lambda_{t,t+1}^U])$	0.04070	(0.01852 , 0.10575)

the stochastic steady state of the model and calculate $(E_t[\Lambda_{t,t+1}], \sigma_t(\Lambda_{t,t+1}))$ at that point. Averaging over these draws yields the black circle. We then undertake a similar procedure but fix the innovation to volatility, $\epsilon_{2,t}$ to be $+1sd$ (green triangle up) and $-1sd$ (red triangle down). We observe substantial variation with innovations to the volatility innovation process. A high volatility shock implies a higher conditional market price of risk to the extent of approximately 15%. The opposite is true in the low volatility case. Table 4 reiterates these results.[15]

Table 5 reports summary statistics of the various components that make up the market price of risk. The moments were computed through simulation using draws from our posterior distribution. We see the standard deviation of the conditional expectation of the stochastic discount factor, $\sigma(E_t[\Lambda_{t,t+1}])$ is low (0.00083). The standard deviation of the risk free rate is also not very volatile. In contrast, the standard deviation of the conditional variance of that stochastic discount factor, $\sigma(V_t[\Lambda_{t,t+1}])$ is two orders of magnitude higher. This means that the movement in the conditional price of risk comes primarily through the movement in the conditional variance term, which, as demonstrated above, comes from movements in the v_t sequence. Importantly, the standard deviation of the conditional variance of $\Lambda_{t,t+1}^U$ is nearly the same as that of of the conditional variance of the entire stochastic discount factor. Recall that this term is much larger than the standard deviation in the conditional expectation. Thus movements in the market price of risk are driven primarily by the $\Lambda_{t,t+1}^U$ term and not $\Lambda_{t,t+1}^R$. Under the doubts interpretation, $\Lambda_{t,t+1}^U$ $(= m_{t+1}(\epsilon_{t+1}))$ is a likelihood ratio that encodes the worst case conditional distribution, relative to the benchmark. Therefore, movements in the conditional market price of risk primarily reflect changes in the distorted conditional

[15] We performed the same exercise but with the state configurations varying with $\epsilon_{1,t}$, at zero and $\pm 1sd$. The market price of risk does not appear to depend on the value of the endowment innovation, which is intuitive given the random walk nature of the (log) endowment process and the IES being fixed at 1.

Table 6: Distorted Moments of the Innovations

v_t	$\mu(\epsilon_1)$	$\mu(\epsilon_2)$	$\sigma^2(\epsilon_1)$	$\sigma^2(\epsilon_2)$	$Cov(\epsilon_1, \epsilon_2)$
High	-0.3904	0.0898	1.0027	1.0126	-0.0605
Med	-0.2233	0.0637	1.0056	1.0068	-0.0361
Low	-0.1370	0.0605	1.0009	1.0068	-0.0157

distribution of the innovations $(\epsilon_{1,t}, \epsilon_{2,t})$. Thus, it is actually concerns for model *uncertainty* that are driving variation in what is typically termed the market price of *risk*.

5.3 Worst Case Distribution

We have established that the interaction of model uncertainty with stochastic volatility has important implications for asset pricing, as captured by the importance of the component $\Lambda^U_{t,t+1}$ in the stochastic discount factor. As discussed above, this likelihood ratio captures the state dependent distortion of the distributions from which the endowment and volatility innovations are drawn. We also showed that this object depended on the level of volatility in the economy. We examine distortions in the context of the lagged value of the volatility component, v_{t-1} being at its non-stochastic steady state and $\pm 2sd$ (where the standard deviations are unconditional).[16] Thus we examine doubts in situations of high, low and moderate variability.

Table 6 displays information on the conditional effects of the evil agent's distortions. We observe clearly the evil agent chooses to lower the mean of the endowment innovation, $\mu(\epsilon_{1,t})$, and that the distortion is more extreme the higher is the volatility state. The second column features distortions in the mean shift of the volatility innovation, $\mu(\epsilon_{2,t})$. Given the undesirability of volatility, this latter distortion involves a positive mean shift. The evil agent shifts the mean of ϵ_2 upwards, with the size of this shift increasing in the level of volatility. The fact that the size of the positive mean shift in ϵ_2 is increasing in the level of volatility means that the evil agent effectively increases the persistence of the perceived volatility process. The evidence regarding the effect on diagonal terms in the covariance matrix is less clear as these are subject to sampling variability.

When one examines the covariances of the shocks under the distorted distributions there appears to be stronger evidence of distorted higher moments. In the high volatility case the

[16]Note that the earlier figure 5 was based on evaluations under deviations in ϵ_2, rather than v_{t-1}. This reflects that fact that the state representation used has v_t decomposed into its lag and innovation components.

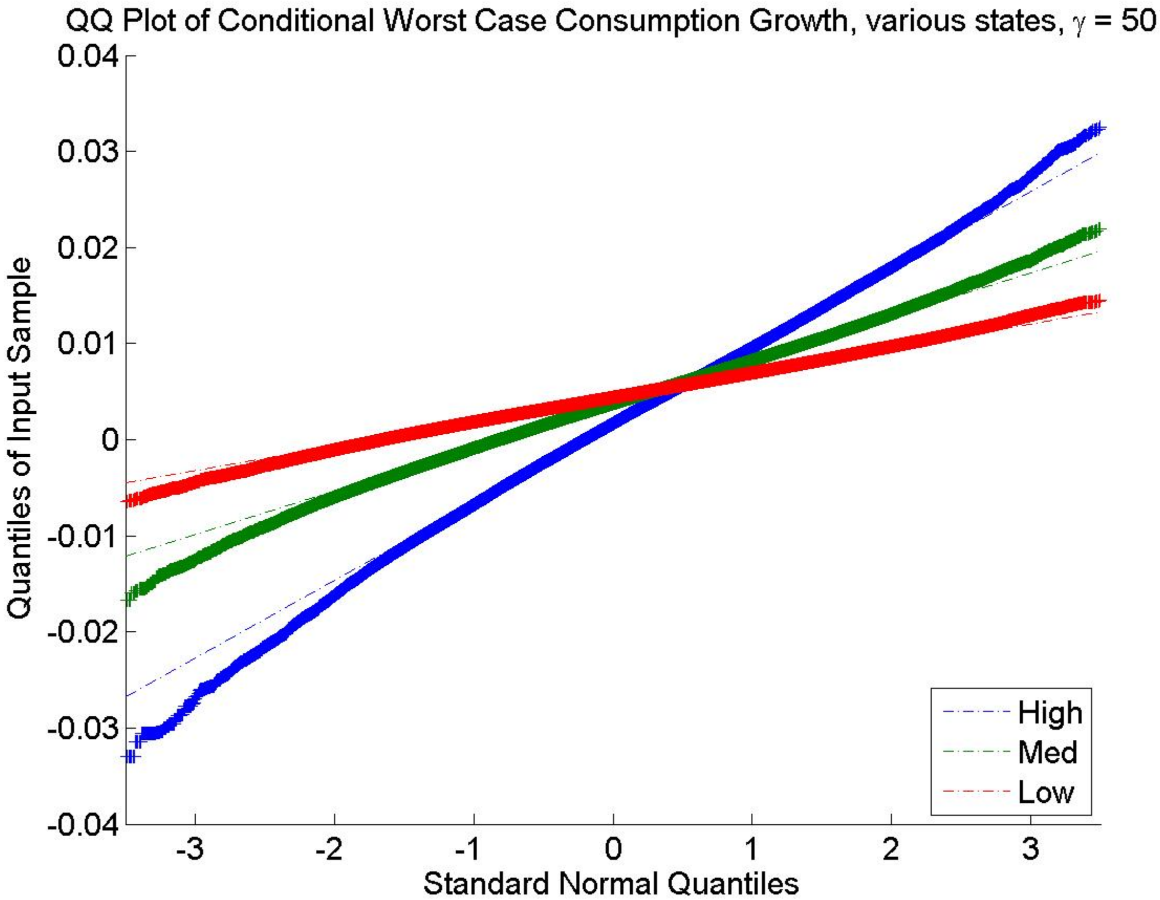


Figure 6: QQ plot of Conditional Worst Case Consumption Growth

evil agent induces a correlation coefficient between the two innovations of approximately 0.06 in absolute value. The correlation is negative, implying that 'good' ('bad') endowment shocks are associated with 'good' ('bad') volatility shocks. Thus, the evil agent makes the good times better and the bad times worse by creating additional systemic risk in the mind of our robust agent. Again we observe that the extent of the distortion is positively related to the current level of volatility as the evil agent takes advantage of a deterioration in detectability.

How do these state dependent distortions of the distribution of the error terms affect the resulting worst case consumption growth process? Figure 6 plots the worst case consumption growth process at the same high, medium, and low volatility states for $\gamma = 50$ using the distorted error terms which made the previous table. We see the distorted means shifting the mean of the worst case consumption growth process. Further, with higher volatility we see a fattening of the left tail of the distribution as evidenced by the departure from the straight line. This departure from normality is more pronounced in the left tail than the right tail. Thus our agent prices assets as if his consumption, conditional in high volatility states, is subject to large decreases.

To examine the implications of the state dependent distortions discussed above for the unconditional distribution of consumption growth under the worst case, we simulate the model under the distorted probability distribution and compute mean, variance, skewness, and autocorrelation. The result of these simulations, for various values of γ are listed in table 7. We observe that by increasing γ (or lowering the penalty of model distortions for the evil agent) the mean of the consumption growth process falls. This is unsurprising, given the consistently negative sign of the distortion to the mean of the innovation to the endowment innovations at various values of the state. Recalling the consistently positive distortion to the mean of the volatility innovation, for various values of the state, it is also unsurprising that the worst case consumption process should be more volatile than in the approximating model. Part of this additional volatility in the endowment sequence also reflects the state dependence of the mean shifts discusses above and the increased persistence in the volatility process under the worst case. Importantly, we find evidence of persistence in the consumption growth sequence. This autocorrelation is small and hard to detect at lower values of γ, but can be seen clearly as γ rises above 35. This stems from the loading of the mean distortion of the endowment innovations on the persistent volatility process, v_t. Consequently, the agent fears a consumption process with a long run risk component. Since the agent prices assets as if using the worst case measure, we therefore can gain some of the asset pricing benefits of long run risk, without taking a stance on whether or not it exists. What is important is that it is thought by the agent to be worth guarding against.

Due to the fact that the innovations $(\epsilon_{1,t}, \epsilon_{2,t})$ are negatively correlated under the distorted model and larger conditional variances are associated with larger (negative) mean shifts in the consumption growth process, we observe negative skewness in the worst case consumption growth process. Since the consumption growth process is persistent, consumption occasionally exhibits swings away from its mean and, due to the nature of the distortions discussed above, the negative swings are more substantial. Much like Rietz (1988) who allows for very infrequent, large drops in consumption growth that have not occurred in postwar consumption data, our model works toward solving the equity premium puzzle by the perceived possibility of disasters in economic growth in the agent's head.

While the induced long run risk component in the agent's worst case consumption growth process is easy to identify from the autocorrelations presented in table 7, it is less clear how

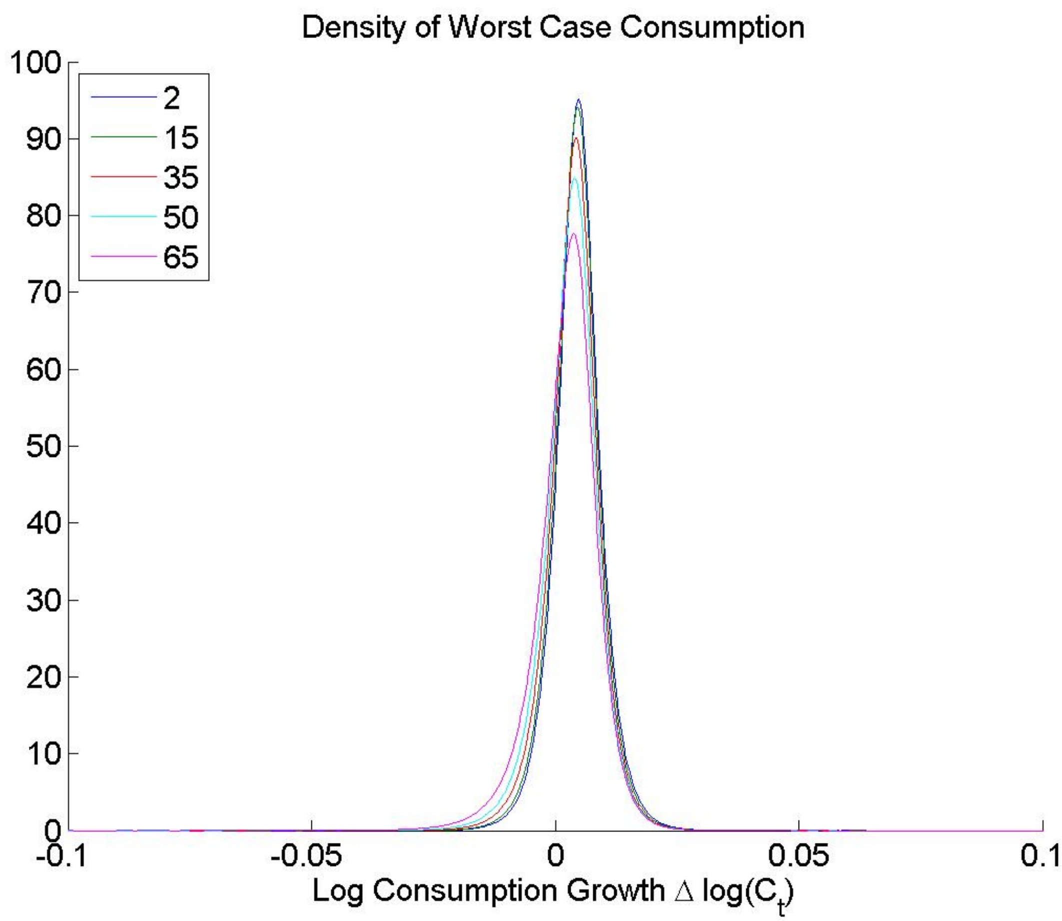


Figure 7: Density of Worst Case Consumption Growth

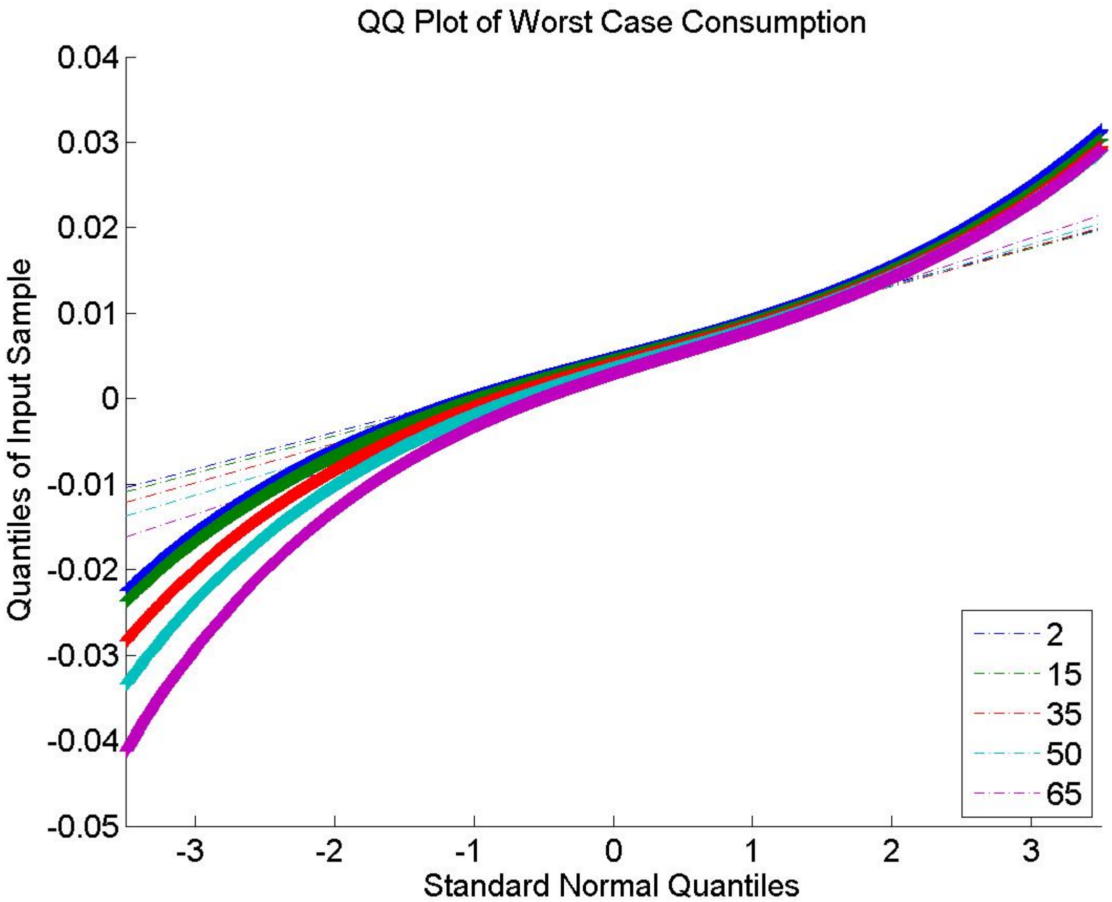


Figure 8: QQ plot of Worst Case Consumption Growth

Table 7: Unconditional Moments of Consumption Under Worst Case Model

γ	2	15	35	50	65
Mean	0.0047	0.0044	0.0038	0.0032	0.0023
St.Dev	0.0051	0.0052	0.0055	0.0059	0.0065
Skew	-0.0145	-0.1157	-0.3100	-0.4684	-0.6598
Corr(lag)					
1	-0.0037	-0.0021	0.0122	0.0321	0.0654
2	-0.0038	-0.0021	0.0102	0.0266	0.0542
3	-0.0038	-0.0029	0.0068	0.0212	0.0463
4	-0.0042	-0.0029	0.0068	0.0184	0.0398
5	-0.0039	-0.0031	0.0045	0.0152	0.0339
6	-0.0033	-0.0029	0.0031	0.0122	0.0276
7	-0.0044	-0.0048	0.0029	0.0105	0.0233
8	-0.0043	-0.0028	0.0001	0.0097	0.0188
9	-0.0036	-0.0041	0.0014	0.0061	0.0148

the skewness reported in that same table induces disasters in our agent's head. To address this issue, figures 7 and 8 present smoothed histograms and qqplots respectively of the agent's worst case consumption growth process for varying levels of γ. In figure 7 we see the slight shift in the mean of the distributions as γ rises. The increase in variance reported in table 7 seems to predominantly come from fattening of the left tail of the distribution with little distortion observable in the right tail. Figure 8 clearly displays both the fat tails induced by our stochastic volatility model relative to the normal distribution and how this fattening (especially of the left tail) increases with γ.

While the induced long run risk component in the agent's worst case consumption growth process is easy to identify from the autocorrelations presented in table 7, it is less clear how the skewness reported in that same table induces disasters in our agent's head. To address this issue, figures 7 and 8 present smoothed histograms and qqplots respectively of the agent's worst case consumption growth process for various levels of γ. In figure 7 we see the slight shift in the mean of the distributions as γ rises. The increase in variance reported in table 7 seems predominantly to come from a fattening of the left tail of the distribution with little distortion observable in the right tail. Figure 8 clearly displays both the fat tails induced by our stochastic volatility model relative to the normal distribution and how this fattening (especially of the left tail) increases with γ.

To examine how frequently 'disasters' happen under the worst case model, we first need

Table 8: Frequency of Disasters (Years)

γ	$\Delta Log(C_t) < -0.02$	$\Delta Log(C_t) < -0.03$
2	570.2	7439.8
15	433.9	4258.6
35	183.7	1403.4
50	89.5	588.7
65	39.2	205.7

define clearly what we mean by a disaster. We will consider situations where consumption falls in a given quarter by more than 2% as a disaster. In table 8 we compute an estimate of the expected waiting time of such disasters and report the number in years.[17] In a nearly undistorted model ($\gamma = 2$) we see that disasters in consumption almost never happen. However, for levels of γ that reach the Hansen-Jaganathan bounds ($\gamma = 35$) disasters happen approximately three times as often. As γ rises, these rare events become much less rare relative to the undistorted model. We should note that these numbers do not take into account the duration of a disaster period, nor whether multiple disasters happen in near consecutive time periods. Since volatility is persistent, and disasters happen precisely when volatility is high, it is conceivable that multiple disasters hit in consecutive or nearly consecutive time periods in which case we should alter our interpretation of the results to take account of this correlation.

5.4 Detectability

Given the important effects of γ on the market price of risk and the worst case probability model, we now assess what is a plausible value of γ, using detection error probabilities. If the two models have similar stochastic properties, they will be difficult to detect using sample sizes that are typically available for analysis. In this case the detection error probability will be close to 0.5, indicating that the models are almost indistinguishable. Models that have very distinguishable characteristics will be easily identifiable and imply a detection error probability of close to 0. We generate data under the true and worst case probability models for γ ranging from 2 to 64. For each γ value, we generate 1000 time series of length 247 under both the true and distorted probability models, and form likelihood ratio test statistics to determine which model generated the data, using the methodology described in

[17] To do this we invert the fraction of the simulated quarters which consumption dropped by that amount, and scale by 1/4.

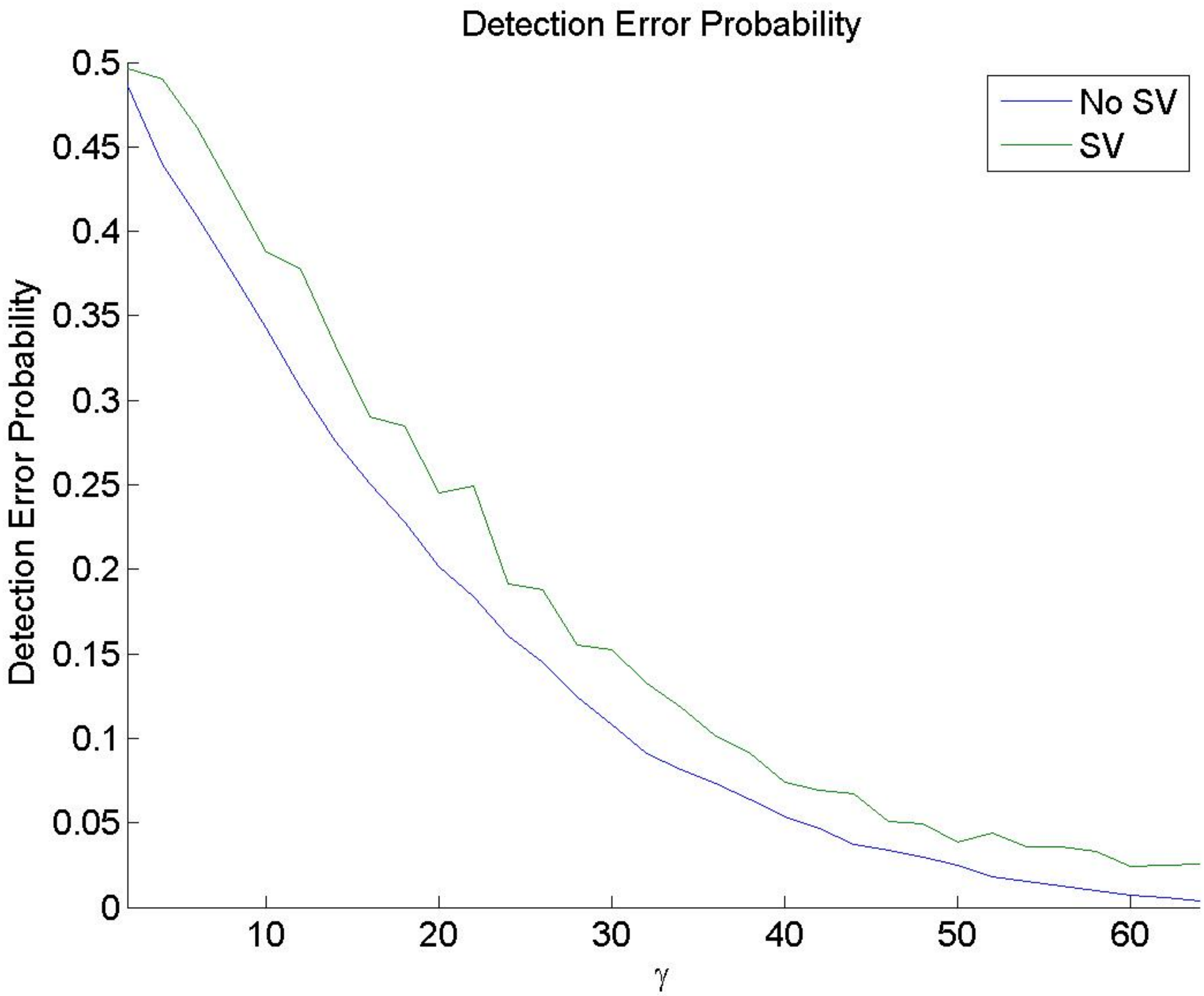


Figure 9: Detection Error Probabilities

the previous section.[18] We compare the detectability with that of the homoskedastic model. All simulations were performed using the same maximum a posteriori parameter values which generated the distorted distributions in section 5.3.

Figure 9 shows that allowing for stochastic volatility in estimation leads to estimates that imply a decreased ability of the agent to discriminate between probability models, for a given value of γ. BHS are able to attain the Hansen-Jagannathan bounds with a γ of about 50, which corresponds to a detection error probability of 2.5%. In our stochastic volatility model, we are able to reach the Hansen-Jagannathan bounds with a $\gamma = 35$. When framed in terms

[18]The particle filters were run with 40000 particles. In cases where there was a divergence in the particle filter, we used likelihood computed up until that time period.

of detection error probabilities, this corresponds to a detection error probability of around 11%.[19] In sections 5.1 and 5.2, we compared the properties of the market price of risk under the heteroskedastic and homoskedastic models for a given level of the penalty parameter (a function of γ). Instead, we can now hold the detection error probability fixed at a level which one finds comfortable. Relative to the homoskedastic model, for a given detection error probability, the stochastic volatility model allows for a larger γ parameter. As we have seen, this helps bring our stochastic discount factor closer to the Hansen-Jaganathan bound.

6 Conclusion

We have illustrated interesting interactions between fears of model misspecification and stochastic volatility, in the context of a simple endowment economy. We show that estimates using US data imply that the model is able to generate a higher market price of risk than estimates in a homoskedastic framework and that there is considerable time variation in the market price of risk, which is absent under homoskedasticity. The asset pricing implications are primarily driven by the component of the stochastic discount factor associated with multiplier preferences. This component allows us to interpret the source of the asset pricing success of the model in terms of a worst case model entertained by the agent. The agent fears lower growth and elevated and more persistent volatility. Most importantly, the agent fears a scenario that features long run risk and occasional disasters in consumption growth. In order to characterize these fears we used a novel set of techniques to draw from the distorted distributions considered by the agent, based on Monte Carlo and perturbation methods that can be applied to a broader class of nonlinear and non-Gaussian models than has previously been considered in the robust control literature.

[19] The detection error probabilities for $\gamma = 34$ and $\gamma = 36$ were 11.8% and 10.1% respectively.

A Appendix

A.1 Perturbation Solution

The (unscaled) risk sensitive value function is:

$$V(C_t, v_t) = (1-\beta)log(C_t) + \frac{\beta}{1-\gamma}log\left(E[(exp(V(C_{t+1}, v_{t+1})))^{(1-\gamma)}]\right)$$

Where:

$$\begin{array}{rcl} C_t & = & exp(Z_t) \\ Z_t & = & \phi + Z_{t-1} + \sigma exp(v_t)\epsilon_{1,t} \\ v_t & = & \lambda v_t + \tau\epsilon_{2,t} \\ \epsilon_t & \sim & N(0, I) \end{array}$$

If we define $\tilde{V}(\epsilon_{1,t}, v_t) = V(C_t, v_t) - log(C_{t-1})$, the detrended value function is:

$$\tilde{V}(\epsilon_{1,t}, v_t) = (1-\beta)(\phi+\sigma exp(v_t)\epsilon_{1,t}) + \frac{\beta}{1-\gamma}log\left(E_t[(exp(\tilde{V}(\epsilon_{t+1}, v_{t+1}))exp(\phi + \sigma exp(v_t)\epsilon_{1,t}))^{(1-\gamma)}]\right)$$

Alternatively, since $\epsilon_{1,t}$ and v_t are known at date t, we can write the scaled value function as

$$\tilde{V}(\epsilon_{1,t}, v_t) \;=\; (\phi + \sigma exp(v_t)\epsilon_{1,t}) + \tfrac{\beta}{1-\gamma}log\left(E_t[(exp(\tilde{V}(\epsilon_{1,t+1}, v_{t+1})))^{(1-\gamma)}]\right)$$

For exposition, we can write the state, $x_t = (\epsilon_{1,t}, v_t) = (\epsilon_t, v_{t-1})$, since we can compute v_t from v_{t-1} and $\epsilon_{2,t}$. We write our minimizing martingale increment as a time invariant function of the state:

$$m_t(\epsilon_t) = m(x_t) = m(\epsilon_{1,t}, \epsilon_{2,t}; v_{t-1})$$

We approximate the log of the minimizing martingale increment by perturbation, so

$$log(m(x_t)) \;\approx\; lm_{ss} + lm_{i,ss}x_t^i + \tfrac{1}{2}lm_{ij,ss}x_t^i x_t^j + \tfrac{1}{6}lm_{ijk,ss}x_t^i x_t^j x_t^k$$

Where lm is short for $log(m(x_t))$. The worst case pdf in our stochastic volatility model is then given by:

$$\tilde{p}(\epsilon_{1,t}, \epsilon_{2,t}; v_{t-1}) \;\approx\; \tfrac{1}{2\pi}exp(lm_{ss} + lm_{i,ss}x_t^i + \tfrac{1}{2}lm_{ij,ss}x_t^i x_t^j + \tfrac{1}{6}lm_{ijk,ss}x_t^i x_t^j x_t^k - \tfrac{1}{2}(\epsilon_{1,t}^2 + \epsilon_{2,t}^2))$$

We show in the main body of text how to use this approximate pdf. We can also define other functional objects to be approximated by taylor polynomials, specifically we can define the conditional expectation and variance of the stochastic discount factor as

$$\begin{array}{rcl} EtL & = & E_t[\Lambda_{t,t+1}] \\ VtL & = & E_t[(\Lambda_{t,t+1} - EtL)^2] \end{array}$$

And approximate these as taylor polynomials. We resort to the law of iterated expectations to estimate the mean of the stochastic discount factor in our calculations $E[E_t[\Lambda_{t,t+1}]] = E[\Lambda_{t,t+1}]$, as well as rely on EtL and VtL to compute variation in the conditional moments and conditional market price of risk as $MPR_t = \frac{\sqrt{VtL}}{EtL}$.

A.2 Accuracy Check

A.2.1 Distorted Distributions

Our method is subject to two sources of approximation error. The first is that we use an approximation to the minimizing Martingale increment and the second is that we are employing Monte Carlo methods. Since the distorted distribution in the homoskedastic model is known exactly we check our method against it. In this context the worst case distortion is given by

$$\tilde{p}(\epsilon) \sim N\left(\mu(\epsilon), 1\right) \tag{14}$$

where

$$\mu(\epsilon) = -\frac{\sigma}{(1-\beta)\theta} \tag{15}$$

Thus, the distortion takes the form of a state independent mean shift, where the size of the shift is increasing in the level of risk in the approximating model and decreasing in θ. The intuition behind the positive effect of σ is that greater variance in the shocks implies that a broader set of distorted models are plausible.

A.2.2 Clouds

In figures 10 and 11 we plot both the posteriors and means for the homoskedastic case with multiplier preferences under various values of gamma, calculated first using closed form expressions and second, using our simulation/perturbation methodology. While the approximate clouds have more dispersion than their closed form counterparts, when we average over the parameter draws, we attain essentially the same values of the market price of risk for each value of γ.

A.3 Computing the Bayes Factor

Suppose we want to compare model 0 and model 1. Want to compute the Bayes Factor

$$\begin{aligned} B_{01} &= \frac{\int p(Y_{1:T}|\vartheta_0)p_0(\vartheta_0)d\vartheta_0}{\int p(Y_{1:T}|\vartheta_1)p_1(\vartheta_1)d\vartheta_1} \\ &= \frac{E_{p_0}[p(Y_{1:T}|\vartheta_0)]}{E_{p_1}[p(Y_{1:T}|\vartheta_1)]} \end{aligned}$$

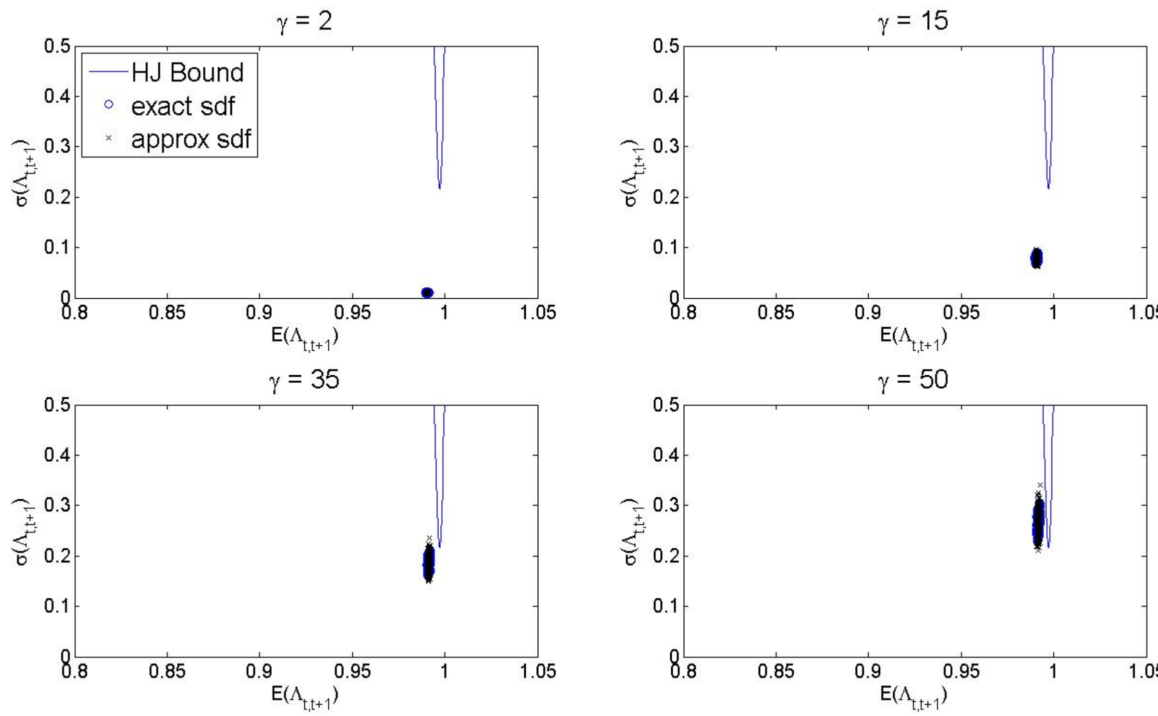


Figure 10: Exact vs Approximated

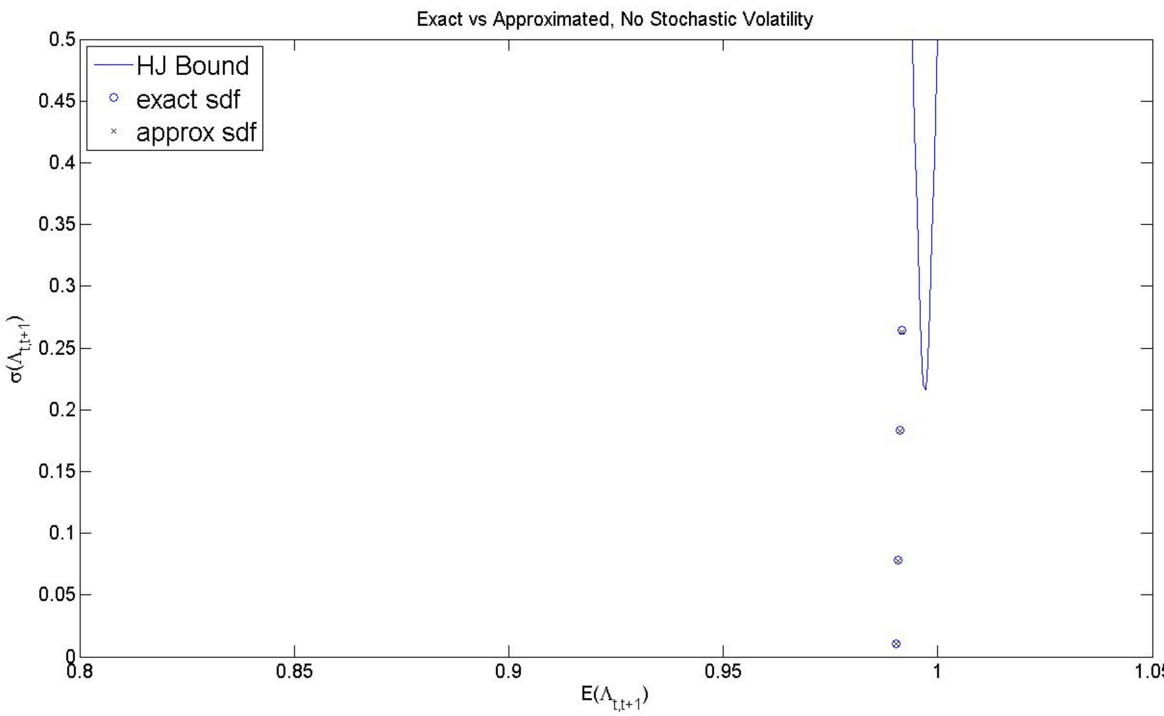


Figure 11: Means: Exact vs Approximated

We compute the marginal likelihood, $E_{p_i}[p(Y_{1:T}|\vartheta_i)] = \int p(Y_{1:T}|\vartheta_i)p_i(\vartheta_i)d\vartheta_i$ for each model i using a harmonic mean estimator. That is, we can estimate,

$$\frac{1}{E_{p_i}[p(Y_{1:T}|\vartheta_i)]} = \int \frac{f_i(\vartheta_i)}{p(Y_{1:T}|\vartheta_i)p_i(\vartheta_i)} p(\vartheta_i|Y_{1:T})d\vartheta_i$$

Where $f_i(\vartheta_i)$ is a density with thinner tails than $p(\vartheta_i|Y_{1:T})$. Specifically, for L draws $\vartheta_i \sim p(\vartheta_i|Y_{1:T})$, we compute

$$\widehat{\vartheta}_i = \frac{1}{L}\sum_{j=1}^{L}\vartheta_i^j$$

and

$$\widehat{\Sigma_{\vartheta_i}} = \frac{1}{L}\sum_{j=1}^{L}\vartheta_i^j$$

Define the support of $f_i(\vartheta_i)$ to be

$$\Theta_i(p) = \left\{\vartheta|(\vartheta - \widehat{\vartheta}_i)'\widehat{\Sigma_{\vartheta_i}}^{-1}(\vartheta - \widehat{\vartheta}_i) < \chi^2_{1-p}(k_i)\right\}$$

Where k_i is the dimension of ϑ_i. Finally we define $f_i(\vartheta_i)$ as

$$f_i(\vartheta_i) = \frac{1}{p}\frac{1}{(2\pi)^{k_i/2}\left|\Sigma_{\vartheta_i}^{1/2}\right|}exp(\frac{-1}{2}(\vartheta - \widehat{\vartheta}_i)'\widehat{\Sigma_{\vartheta_i}}^{-1}(\vartheta - \widehat{\vartheta}_i))I_{\Theta_i(p)}(\vartheta_i)$$

References

Anderson, E. W., 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.

Andrieu, C., A. Doucet, and R. Holenstein (2010): "Particle Markov Chain Monte Carlo with discussion," *Journal of Royal Statistical Society Series B*, 72, 269 – 342.

Aruoba, W., J. Fernández-Villaverde, and J. F. Rubio-Ramírez (2006): "Comparing Solution Methods for Dynamic Equilibrium Economies," *Journal of Economic Dynamics and Control*, 30, 2477–2508.

Bansal, R., and A. Yaron (2004): "Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59(4), 1481–1509.

Barillas, F., L. P. Hansen, and T. J. Sargent (2009): "Doubts or Variability," *Journal of Economic Theory*, 144, 2388–2418.

Barro, R. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*.

Barro, R., and J. F. Ursua (2008a): "Consumption Disasters in the Twentieth Century," *American Economic Review, Papers and Proceedings*, 98, 58–63.

——— (2008b): "Macroeconomic Crisis since 1870," *Brooking Papers on Economic Activity*, Spring, 255–350.

Bloom, N., M. Floetotto, and N. Jaimovich (2009): "Really Uncertain Business Cycles," Working paper, Stanford University.

Clark, T. E. (2009): "Is the Great Moderation Over? An empirical investigation," *Economic Review, Federal Reserve Board of Kansas City*, QIV, 5–42.

Creal, D. (2010): "A survey of sequential Monte Carlo methods for economics and finance," Econometric Reviews, forthcoming.

Drechsler, I. (2010): “Uncertainty, Time-Varying Fear, and Asset Prices,” Working Paper, New York University.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez (2007): “Estimating Macroeconomic Models: A Likelihood Approach,” *Review of Economic Studies*, 74, 1059–1087.

Flury, T., and N. Shephard (2008): “Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models,” Unpublished paper, University of Oxford.

Hansen, L. P., and R. Jagannathan (1991): “Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns,” *Journal of Political Economy*, 99, 225–262.

Hansen, L. P., and T. J. Sargent (2008): *Robustness*. Princeton University Press, New York, NY, 1 edn.

Hansen, L. P., and K. Singleton (1982): “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models,” *Econometrica*, 50(5), 1269–86.

Justiniano, A., and G. Primiceri (2008): “The Time Varying Volatility of Macroeconomic Fluctuations,” *American Economic Review*, 98, 604–641.

Kleshchelski, I., and N. Vincent (2008): “Robust Equilibrium Yield Curves,” Working Paper, HEC Montreal.

Lucas, R. (1978): “Asset Prices in an Exchange Economy,” *Econometrica*, 46, 1429–1446.

McConnell, M. M., and G. Perez-Quiros (2000): “Output Fluctuations in the United States: What Has Changed Since the Early 1980’s,” *American Economic Review*, 90, 1464–1476.

Mehra, R., and E. C. Prescott (1985): “The Equity Premium Puzzle,” *Journal of Monetary Economics*, 15, 145–161.

Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller (1953): “Equation of State Calculations by Fast Computing Machines,” *Journal of Chemical Physics*, 21, 1087 – 1092.

Rietz, T. A. (1988): "The Equity Risk Premium: A Solution?," *Journal of Monetary Economics*, 22(1), 117–131.

Rubin, D. B. (1987): "A noniterative sampling importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: the SIR algorithm," *Journal of the American Statistical Association*, 52, 543–546.

Smith, A., and A. Gelfand (1992): "Bayesian Statistics Without Tears: A Sampling-Resampling Perspective," *The American Statistician*, 46, 84–88.

Smith, M. (2010): "Estimating Nonlinear Economic Models Using Surrogate Transitions," Working paper, New York University.

Stock, J., and M. Watson (2002): "Has the Business Cycle Changed and Why," NBER Working Paper Series.

Tallarini, T. D. (2000): "Risk-Sensitive Real Business Cycle Models," *Journal of Monetary Economics*, 45, 507–532.

Ursua, J. F. (2010): "Long Run Volatility," Working Paper, Harvard University.