Fragile Beliefs and the Price of Uncertainty

Lars Peter Hansen*

Thomas J. Sargent‡

December 26, 2008

Abstract

A representative consumer uses Bayes' law to learn about parameters and to construct probabilities with which to perform ongoing model averaging. The arrival of signals induces the consumer to alter his posterior distribution over parameters and models. The consumer copes with specification doubts by slanting probabilities pessimistically. One of his models puts long-run risks in consumption growth. The pessimistic probabilities slant toward this model and contribute a counter-cyclical and signal-history-dependent component to prices of risk.

Key words: Learning, Bayes' law, robustness, risk-sensitivity, pessimism, prices of risk.

---

*University of Chicago. Email: l-hansen@uchicago.edu.

†New York University and Hoover Institution. Email: ts43@nyu.edu.

‡We thank Gadi Barlevy, Alberto Bisin, Riccardo Colacito, Mark Gertler, Anastasios Karantounias, Ricardo Mayer, Tomasz Piskorski, Grace Tsiang, and Gianluca Violante for helpful comments on earlier drafts of this paper. We thank Francisco Barillas, Ricardo Mayer, and Leandro Nascimento for excellence in executing the computations. We thank the National Science Foundation for research support under separate grants to Hansen and Sargent.

Le doute n'est pas une condition agréable, mais la certitude est absurde.[1]
Voltaire 1767.

1 Introduction

A pessimist thinks that good news is temporary and that bad news endures. This paper describes how a representative consumer's model selection problem and fear of model misspecification foster pessimism that puts countercyclical model uncertainty premia into risk prices.

1.1 How doubts promote fragile beliefs

Our representative consumer values consumption streams according to iterated versions of the multiplier preferences that Hansen and Sargent (2001) use to represent aversion to model uncertainty.[2] Following Hansen and Sargent (2007), an iterated application of risk-sensitivity operators allows us to focus the representative consumer's ambiguity on particular aspects, including model selection and parameter values.[3] *Ex post*, the consumer acts 'as if' he uses a probability measure that he twists pessimistically relative to his approximating model. By 'fragile beliefs' we refer to the responsiveness of pessimistic probabilities to the arrival of news, as determined by the state dependent value functions that define what the consumer is pessimistic *about*.[4] Relative to the conventional rational expectations case in which the representative consumer has complete confidence in his statistical model, our representative consumer's reluctance fully to trust a single approximating model adds 'model uncertainty premia' to prices of risk. New uncertainty components of the "risk" prices emerge from the hidden Markov model that we impute to the consumer. They are time-dependent and state-dependent, in contrast to the constant uncertainty premium analyzed by Hansen et al. (1999) and Anderson et al. (2003).

[1]Doubt is not a pleasant condition, but certainty is absurd.

[2]The relationship of the multiplier preferences of Hansen and Sargent (2001) to the max-min expected utility preferences of Gilboa and Schmeidler (1989) are analyzed by Hansen et al. (2006), Maccheroni et al. (2006a,b), Cerreia et al. (2008), and Strzalecki (2008).

[3]Sometimes the literature calls this 'structured uncertainty'.

[4]Harrsion and Kreps (1978) and Scheinkman and Xiong (2003) explore another setting in which difficult to detect departures from rational expectations lead to interesting asset price dynamics that cannot occur under rational expectations.

1.2 Fragile expectations as sources of time-varying risk premia

A hidden Markov model for consumption growth confronts a representative consumer with ongoing model selection and parameter estimation problems. Our representative consumer wants to know components of a hidden state vector, some that stand for unknown parameters within a model and others that index models. A probability distribution over that hidden state becomes part of the state vector in the representative consumer's value function. Bayes' law describes its motion over time. The representative consumer slants probabilities towards the model that has the lowest utility. We show how variations over time in the probabilities attached to models and other state variables put volatility into the model uncertainty premia.

1.3 Key components

In addition to the risk sensitivity operator that Tallarini (2000) applied, we introduce an additional one, taken from Hansen and Sargent (2007), that adjusts the probability distribution of hidden Markov states for model uncertainty.[5] We interpret both risk-sensitivity operators as capturing the representative consumer's concerns about robustness instead of the enhanced risk aversion interpretation of Tallarini.[6]

Our representative consumer assigns positive probabilities to two models whose fits make them indistinguishable for our data on per capita U.S. consumption expenditures on nondurables and services from 1948II-2008III. In one model, consumption growth rates are nearly i.i.d. model, and in the other there is a highly persistent component to the consumption growth rate as in the long-run risk model of Bansal and Yaron (2004) with a persistent component in consumption growth. But the consumer doubts the model-mixing probabilities as well as the specification of each of the component models. In contrast, Bansal and Yaron assume that the representative consumer assigns probability one to the long-run risk model even though sample evidence is indecisive in selecting between them.[7] Our framework explains why a

[5] This second risk-sensitivity operator accounts for what some researchers call 'structured' model uncertainty.

[6] Barillas et al. (2007) reinterpret some of Tallarini's results in terms of concern about model misspecification instead of risk aversion.

[7] Bansal and Yaron (2004) incorporate other features in their specifications of consumption dynamics, including stochastic volatility. They also use a recursive utility specification with an in-

consumer might act *as if* he puts probability (close to) one on the long-run risk model even though he knows that it is difficult to discriminate between these models statistically.

1.4 Relation to other asset pricing models with Bayesian learning

By adding model uncertainty, we exploit substantially more of the hidden state structure than did earlier researchers, for example, Detemple (1986), David (1997), Veronesi (2000), Brennan and Xia (2001), Ai (2006), and Croce et al. (2006),[8] who use learning about a hidden state simply to generate an exogenous process for distributions of future signals conditional on past signals as an input into a consumption based asset pricing model. Those papers specify the evolution of a primitive stochastic process of technology or endowments with an information structure that conceals hidden Markov states. By applying Bayesian learning, typically as embodied in recursive filtering methods, they construct a less informative state vector that consists of sufficient statistics for the distribution of hidden states conditioned on signal histories, as well as a recursive law of motion for it that is then used as an input in decision making and asset pricing. Having constructed the coarser information structure implied by Bayesian learning, decision making and asset pricing in these models is standard. Therefore, the asset pricing implications of such learning models depend only on the distributions of future signals conditioned on past signals, and not on the underlying structure with hidden states that the model builder used to deduce that distribution. In such models, the only thing that learning contributes is a justification for those conditional distributions: we would get equivalent asset pricing implications by just *assuming* those distributions from the start.

As we shall see, application of a risk-sensitivity operator to twist the distribution of hidden states means that that equivalence is not true in our model because it

tertemporal elasticity of substitution greater than 1.

[8]The learning problems in those papers share the feature that learning is passive, there being no role for experimentation so that prediction can be separated from control. Cogley et al. (2005) apply the framework of Hansen and Sargent (2007) in a setting where decisions affect future probabilities of hidden states and experimentation is active. The papers just cited price risks under the same information structure that is used to generate the risks being priced. In section 5, we offer an interpretation of some other papers (e.g., Bossaerts (2002, 2004) and Cogley and Sargent (2008)) that study the effects of agents Bayesian learning on pricing risks generated by limited information sets from the point of view of an outside econometrician who has a *larger* information set.

makes asset prices depend on the evolution of the hidden states and not simply on the distribution of future signals conditioned on signal histories. This occurs because of how, following Hansen and Sargent (2007), we make the representative consumer explore potential misspecifications of the distributions of hidden Markov states and of future signals conditioned on those hidden Markov states and on how he therefore refuses to reduce compound lotteries. Continuation utilities will be center stage in how our representative consumer uses signal histories to learn about hidden Markov states, an ingredient absent from those earlier applications of Bayesian learning that reduced the representative consumer's information prior to asset pricing.[9]

1.5 Organization

We proceed as follows. After section 2 sets out a framework for pricing risks implicit in a vector Brownian motion w_t, section 3 describes a hidden Markov model and three successively less information structures (full information, unknown states, and unknown states and unknown model) together with the three innovations (or news) processes given by the increments to $W_t(\iota), \bar{W}_t(\iota)$ and $\bar{W}_t$ that are implied by these three information structures. Section 4 then uses these three information specifications and associated choices $dW_t(\iota), d\bar{W}_t(\iota)$ and $d\bar{W}_t$ as the risks dw_t to be priced without model uncertainty. We construct these section 4 risk prices under the information assumptions ordinarily used in finance and macroeconomics. Section 5 offers a different perspective on Bayesian learning by pricing each of the risks $dW_t(\iota), d\bar{W}_t(\iota)$ and $d\bar{W}_t$ under the single full information set. Section 6 describes contributions to risk prices coming from model uncertainties about distributions conditioning on each of our three information sets. Uncertainty about shock distributions with known states contributes a constant uncertainty premium, while uncertainty about unknown states contributes a time-dependent one and uncertainty about models contributes a state-dependent one. Section 7 presents an empirical example designed to highlight the mechanism through which the state-dependent uncertainty premia give rise to countercyclical prices of risk. Appendix A describes how we use detection error

[9]Exceptions to this statement are our earlier papers Cagetti et al. (2002) and Hansen et al. (2002) that incorporate robustness corrections. In a sequel to this paper, Hansen (2007), among other things (a) expands the model uncertainty faced by the representative consumer by effectively confronting him with a continuum of long-run risk models parameterized by Bansal and Yaron's ρ, and (b) studies the consequences of applying a risk-sensitivity adjustment to this additional source of model uncertainty.

probabilities to calibrate the representative consumer's concerns about model misspecification, while appendix B proliferates models as part of a robustness exercise designed to refine our understanding of the forces that produce countercyclical risk prices.

2 Stochastic discounting and risks

Let $\{S_t\}$ be a stochastic discount factor process that, in conjunction with an expectation operator, assigns date 0 risk-adjusted prices to payoffs at date t. Trading at intermediate dates implies that $\frac{S_{t+\tau}}{S_t}$ is the τ-period stochastic discount factor for pricing at date t. Let $\{w_t\}$ be a vector Brownian motion innovation process where the increment dw_t represents new information flowing to consumers at date t. We synthesize a cumulative time t payoff as

$$\log Q_t(\alpha) = \alpha \cdot (w_t - w_0) - \frac{t}{2}|\alpha|^2.$$

By subtracting $\frac{t}{2}|\alpha|^2$, we make the payoff be a martingale with unit expectation. By changing the vector α, we change the risk exposure to components of w_t. At date t, we price the payoff $\frac{Q_{t+\tau}(\alpha)}{Q_t(\alpha)}$ as

$$P_{t,\tau}(\alpha) = E\left[\frac{S_{t+\tau}Q_{t+\tau}(\alpha)}{S_t Q_t(\alpha)}\Big|\mathcal{F}_t\right]. \tag{1}$$

The vector of (growth-rate) risk prices for horizon τ is given by the price "elasticity":

$$\pi_{t,\tau} = -\frac{\partial}{\partial\alpha}\frac{1}{\tau}\log P_{t,\tau}(\alpha)|_{\alpha=\alpha_o}, \tag{2}$$

where we have scaled by the payoff horizon τ for comparability. We take the negative because exposure to risk is bad. Since we scaled the payoffs to have unit price, $-\frac{1}{\tau}\log p_{t,\tau}$ is the logarithm of an expected return adjusted for the payoff horizon. In log-normal models, this derivative is independent of α_o. This is true more generally when the investment horizon shrinks to zero.[10]

[10]Here we are following Hansen and Scheinkman (2009) and Hansen (2008) in constructing a term structure of prices of growth-rate risk.

The vector of local risk prices is given by the limit

$$\pi_t = -\lim_{\tau \downarrow 0} \frac{\partial}{\tau \partial \alpha} \log P_{t,\tau}. \tag{3}$$

It gives the local compensation for exposure to shocks expressed as an increase in the conditional mean return. Local risk prices in conjunction with an instantaneous risk-free rate are the building blocks of asset prices (*e.g.*, Duffie (2001, pp. 111-114)). These local prices can be compounded to construct the asset prices for arbitrary payoff intervals τ using the dynamics of the underlying state variables in an economy.

We exploit the local normality to obtain a simple characterization of the slope of the mean-standard deviation frontier and to reproduce a classical result from finance. The slope of the efficient segment of the mean-standard deviation frontier is obtained by solving:

$$\max_{\alpha, \alpha \cdot \alpha = 1} \alpha \cdot \pi_t$$

where the constraint imposes a unit local variance. The solution is $\alpha_t^* = \frac{\pi_t}{|\pi_t|}$ with the optimized local mean given by

$$\alpha_t^* \cdot \pi_t = \frac{\pi_t \cdot \pi_t}{|\pi_t|} = |\pi_t|, \tag{4}$$

In this local normal environment, the Hansen and Jagannathan (1991) analysis simplifies to comparing the magnitude of the risk price vector implied by alternative models to an observed mean-standard deviation frontier.

In the power utility model,

$$\frac{S_{t+\tau}}{S_t} = \exp(-\delta) \exp[-\gamma(c_{t+\tau} - c_t)],$$

where the growth rate of log consumption $c_{t+\tau} - c_t$. Here the vector π_t of local risk prices is the vector of "exposures" of $-d \log S_t = \gamma dc_t$ to the Brownian increment vector dw_t.

To study learning and robustness, we use models of Bayesian learning to create alternative specifications of w and information sets with respect to which the mathematical expectation in (1) are evaluated.

3 Three information structures

We use a hidden Markov model and two filtering problems to construct three information sets that we shall use to define risks to be priced with and without concerns about robustness to model misspecification.

3.1 State evolution

Two models $\iota = 0, 1$ take the state-space forms

$$\begin{aligned} d\zeta_t(\iota) &= A(\iota)\zeta_t(\iota)dt + B(\iota)dW_t \\ dy_t &= D(\iota)\zeta_t(\iota)dt + G(\iota)dW_t \end{aligned} \tag{5}$$

where $\zeta_t(\iota)$ is the *state*, y_t is the *signal*, and W is a multivariate standard Brownian motion. For notational simplicity, we suppose that the same Brownian motion drives both models. Under full information ι is observed and the vector dW_t gives the new information available to the consumer at date t.

3.2 Filtering problems

To generate two alternative information structures, we solve two types of filtering problem. Let $\mathcal{Y}_t$ be generated by the history of the signal dy_τ up to t. In what follows we first condition on $\mathcal{Y}_t$ and ι for each t. We then omit ι from the consumer's conditioning information.

3.2.1 Model known

First, suppose that ι is known. Application of the Kalman filter yields the following innovations representation:

$$d\bar{\zeta}_t(\iota) = A(\iota)\bar{\zeta}_t(\iota) + K_t(\iota)[ds_t - D(\iota)\bar{\zeta}_t(\iota)]$$

where $\bar{\zeta}_t(\iota) = E[\zeta_t(\iota)|\mathcal{Y}_t, \iota]$ and

$$K_t(\iota) = [B(\iota)G(\iota)' + \Sigma_t(\iota)D(\iota)'][G(\iota)G(\iota)']^{-1}$$

$$
\begin{aligned}
\frac{d\Sigma_t(\iota)}{dt} &= A(\iota)\Sigma_t(\iota) + \Sigma_t A(\iota)' + B(\iota)B(\iota)' \\
&\quad -K_t(\iota)[G(\iota)B(\iota)' + D(\iota)\Sigma_t(\iota)].
\end{aligned} \tag{6}
$$

The *innovation process* is

$$
d\bar{W}_t(\iota) = [\bar{G}(\iota)]^{-1}\left[ds_t - D(\iota)\bar{\zeta}_t(\iota)dt\right]
$$

where $G(\iota)G'(\iota) = \bar{G}(\iota)\bar{G}(\iota)'$ and $\bar{G}(\iota)$ is nonsingular. The innovation process constitutes the new information revealed to economic agents by the signal history.

3.2.2 Model unknown

Assume that $G(\iota)G(\iota)'$ independent of ι. Without this assumption, ι is revealed immediately. Let $\bar{\iota}_t = E(\iota|\mathcal{Y}_t)$ and

$$
d\bar{W}_t = \bar{G}^{-1}\left(ds_t - \nu_t dt\right) = \bar{\iota}_t d\bar{W}_t(1) + (1 - \bar{\iota}_t)d\bar{W}_t(2)
$$

where

$$
\nu_t \doteq [\bar{\iota}_t D(1)\bar{\zeta}_t(1) + (1 - \bar{\iota}_t)D(0)\bar{\zeta}_t(0)]. \tag{7}
$$

Then

$$
d\bar{\iota}_t = \bar{\iota}_t(1 - \bar{\iota}_t)[\bar{\zeta}_t(1)'D(1)' - \bar{\zeta}_t(0)'D(0)']\left(\bar{G}'\right)^{-1} d\bar{W}_t.
$$

The new information pertinent to consumers is now $d\bar{W}_t$.

4 Risk prices

Section 3 described three information structures: i) full information, ii) hidden states but known model, iii) unknown states and unknown model. We use the associated Brownian motions $W(\iota), \bar{W}_t(\iota)$, and $\bar{W}_t$ as risks to be priced and price those risks under the information structure that generated them. The forms of the risk prices are the same for all three information structures and are familiar from Breeden (1979). (But in section 5 we shall price all three risks under full information in order to look at Bayesian learning from another angle.) Given the local normality of the diffusion model, the risk prices are given by the exposures of the log marginal utility to the underlying risks. Let the increment logarithm of consumption be given by

$dc_t = H'dy_t$, implying that consumption growth rates are revealed by the increment in the signal vector. Each of the differing information sets implies a risk price vector, as reported in Table 1.

Because different risks are being priced, the risk prices change across information structures. However, the *magnitude* of the risk price is the same across information structures. As we saw in (4), the magnitude of the risk price vector is the slope of the instantaneous mean-standard deviation frontier. In section 6, we shall show how a concern about model misspecification alters risk prices by adding compensations for bearing model uncertainty. But first we want to look at Bayesian learning and risk prices from a different perspective.

information	local risk	risk price	slope
full	dW_t	$\gamma G(\iota)'H$	$\gamma\sqrt{H'G(\iota)G(\iota)'H}$
unknown state	$d\bar{W}_t(\iota)$	$\gamma\bar{G}(\iota)'H$	$\gamma\sqrt{H'G(\iota)G(\iota)'H}$
unknown model	$d\bar{W}_t$	$\gamma\bar{G}'H$	$\gamma\sqrt{H'G(\iota)G(\iota)'H}$

Table 1: When the model is unknown, $G(\iota)G(\iota)'$ is assumed to be independent of ι. The parameter γ is the coefficient of relative risk aversion in a power utility model. The entries in the "slope" column are the implied slope of the mean-standard deviation frontier. The consumption growth rate is $dc_t = H'dy_t$.

5 A full-information perspective on agents' learning

In this section, we study what happens when an econometrician mistakenly presumes that consumers have a larger information set than they actually do. It is known that an econometrician who conditions on *less* information than consumers nevertheless draws correct inferences about the magnitude of risk prices. But we shall see that an econometrician who mistakenly conditions on *more* information than consumers actually have makes false inferences about that magnitude. We regard the consequences of an econometrician's mistakenly conditioning on more information than consumers as contributing to the analysis of risk pricing under consumers' Bayesian learning.

Hansen and Richard (1987) systematically studied the consequences for risk prices of an econometrician's conditioning on less information than consumers. Given a correctly specified stochastic discount factor process, if economic agents use *more* information than an econometrician, the consequences for the econometrician's inferences about risk prices can be innocuous. In constructing conditional moment restrictions for asset prices, all that is required is that the econometrician at least include prices in his information set. By application of the law of iterated expectation, the product of a cumulative return and a stochastic discount factor remains a martingale when some of the information available to consumers is omitted from the econometrician's information set. While the econometrician who omits information fails correctly to infer the risk components actually confronted by consumers, that mistake does not undermine his correct inference about the slope of the mean-standard deviation frontier, as we saw in the third column of table 1 section 3.

We now consider the reverse problem. What happens if economic agents use *less* information than an econometrician? We study this by using the full-information structure but price risks generated by the smaller informative information structures, in particular, $d\bar{W}_t(\iota)$ and $d\bar{W}_t$. In pricing $d\bar{W}_t(\iota)$ and $d\bar{W}_t$ under full information, we use pricing formulas that take the mistaken Olympian perspective (often used in macroeconomics) that the consumers know the full-information probability distribution of signals. This mistake made by the econometrician induces a pricing error relative to the risk prices that are actually confronted by the consumer. The full information prices misrepresent the "risks" consumers confront with their reduced information structures. The price discrepancies represent the effects of a representative agent's learning that Bossaerts (2002, 2004) and Cogley and Sargent (2008) featured.

5.1 Hidden states but known model

Consider first the case in which the model is known. Represent the innovation process as

$$d\bar{W}_t(\iota) = [\bar{G}(\iota)]^{-1} \left(D(\iota) \left[\zeta_t(\iota) - \bar{\zeta}_t(\iota) \right] dt + G(\iota) dW_t \right).$$

This expression reveals that $d\bar{W}_t(\iota)$ bundles two risks: $\zeta_t - \bar{\zeta}_t$ and dW_t. An innovation under the reduced information structure ceases to be an innovation in the original full information structure. Also, the "risk" $\zeta_t(\iota) - \bar{\zeta}_t(\iota)$ under the limited information structure ceases to be risk under the full information structure.

Consider the pricing of the small time interval limit of

$$\frac{\bar{Q}_{t+\tau}(\bar{\alpha})}{Q_t(\bar{\alpha})} = \exp\left(\bar{\alpha}' \left[\bar{W}_{t+\tau}(\iota) - \bar{W}_t(\iota)\right] - \frac{|\bar{\alpha}|^2 \tau}{2}\right).$$

This has unit expectation under the partial information structure. The local price computed under the full information structure is:

$$-\delta - \gamma H \zeta_t(\iota) + \bar{\alpha}'[\bar{G}(\iota)]^{-1} D(\iota) \left[\zeta_t(\iota) - \bar{\zeta}_t(\iota)\right] + \frac{1}{2}\left|-\gamma H' G(\iota) + \bar{\alpha}' \left[\bar{G}(\iota)\right]^{-1} G(\iota)\right|^2 - \frac{|\bar{\alpha}|^2}{2}$$

where δ is the subjective rate of discount. Multiplying by minus one and differentiating with respect to $\bar{\alpha}$ gives the local price:

$$\gamma \bar{G}(\iota)' H + [\bar{G}(\iota)]^{-1} D(\iota) \left[\bar{\zeta}_t(\iota) - \zeta_t(\iota)\right].$$

The first term is the risk price under partial information (see Table 1), while the second term is the part of the forecast error in the signal under the reduced information set that can be forecast perfectly under the full information set.

5.2 States and model both unknown

Consider next what happens when the model is unknown. Suppose that $\iota = 1$ and represent $\bar{W}_t$ as

$$\bar{W}_t = \bar{G}^{-1}\left[G(1) dW_t + D(1)\zeta_t(1) dt\right] - \bar{G}^{-1}\left[\bar{\iota}_t D(1)\bar{\zeta}_t(1) dt + (1 - \bar{\iota}_t) D(0)\bar{\zeta}_t(0) dt\right]$$

There is an analogous calculation for $\iota = 0$. When we compute local prices under full information, we obtain

$$\gamma \bar{G}' H + \bar{G}^{-1}\left[\nu_t - D(\iota)\zeta_t\right] \tag{8}$$

where ν_t is defined in (7).

The term $\gamma \bar{G}' H$ is the risk price under reduced information when the model is unknown (see Table 1). The term $\bar{G}^{-1}\left[\nu_t - D(\iota)\zeta_t\right]$ is a contribution to the risk price measured by the econometrician coming from the effects of the consumer's learning on the basis of his more limited information set. With respect to the probability distribution used by the consumer, this term averages out to zero. Since ι is unknown, the average includes a contribution from the prior. For some sample paths, this term can have negative entries for a substantial amount of time, indicating that the prices under the reduced information exceed those computed under full information. Other trajectories could display just the opposite phenomenon. It is thus possible that the term $\bar{G}^{-1}[\nu_t - D(\iota)\zeta_t]$ contributes apparent pessimism or optimism, depending on the prior over ι and the particular sample path. In what follows, we use concerns about robustness to motivate priors that are *necessarily* pessimistic and that always enhance the counterpart to risk prices.

6 Price effects of consumers' concerns about robustness

When prices reflect a representative consumer's fears of model misspecification, (2) must be replaced by

$$P_{t,\tau}(\alpha) = \hat{E}\left[\frac{S_{t+\tau}Q_{t+\tau}(\alpha)}{S_tQ_t(\alpha)}\Big|\mathcal{F}_t\right], \tag{9}$$

where $\hat{E}$ is the mathematical expectation with respect to the representative consumer's worst-case probability distribution. This equation can also be expressed as

$$P_{t,\tau}(\alpha) = E\left[\frac{M_{t+\tau}}{M_t}\frac{S_{t+\tau}Q_{t+\tau}(\alpha)}{S_tQ_t(\alpha)}\Big|\mathcal{F}_t\right], \tag{10}$$

where $M_{t+\tau}$ is a martingale that represents the ratio of a worst-case density to the original density for the Brownian motion $w_{t+\tau}$. To capture the effects of a preference for robustness, we first construct the likelihood ratio $\frac{M_{t+\tau}}{M_t}$ for the representative consumer, then proceed to compute associated prices of risk and the implied slopes of the efficient frontier by using formula (4). To compute $\frac{M_{t+\tau}}{M_t}$ under our alternative information structures, we must find value functions for a planner who fears model misspecification.[11] We do this below for our three information structures and then construct the last column of Table 2, which indicates the contribution to risk prices from each type of model uncertainty.

[11]Hansen and Sargent (2008, chs.11-13) discuss the role of the planner's problem in computing and representing prices with which to confront a representative consumer.

information	local risk	risk price	uncertainty price
full	dW_t	$G(\iota)'H$	$\frac{1}{\theta_1}[B(\iota)'\lambda(\iota) + G(\iota)'H]$
unknown state	$d\bar{W}_t(\iota)$	$\bar{G}(\iota)'H$	$\frac{1}{\theta_2}\bar{G}(\iota)^{-1}D(\iota)\Sigma_t(\iota)\lambda(\iota)$
unknown model	$d\bar{W}_t$	$\bar{G}'H$	$(\tilde{\iota} - \bar{\iota})\bar{G}^{-1}[D(1)\bar{\zeta}(1) - D(0)'\bar{\zeta}(0)]$

Table 2: When the model is unknown, $G(\iota)G(\iota)'$ is assumed to be independent of ι. The consumption growth rate is $dc_t = H'ds_t$. Please *cumulate* contributions to uncertainty prices as you move down the last column.

6.1 Value function without robustness

We study a consumer with unitary elasticity of intertemporal substitution. We start with the value function for discounted expected utility using a logarithm period utility function:

$$\begin{aligned}
V(\zeta, c, \iota) &= \delta E\left[\int_0^\infty \exp(-\delta\tau)c_{t+\tau}|\zeta_t = \zeta, c_t = c, \iota\right] \\
&= \delta E\left[\int_0^\infty \exp(-\delta\tau)(c_{t+\tau} - c_t)|\zeta_t = \zeta, c_t = c, \iota\right] + c \\
&= \lambda(\iota)\cdot\zeta + c.
\end{aligned}$$

Given the recursive nature of this valuation, the vector $\lambda(\iota)$ satisfies the equation

$$0 = -\delta\lambda(\iota) + D(\iota)'H + A(\iota)'\lambda(\iota), \tag{11}$$

and thus

$$\lambda(\iota) = [\delta I - A(\iota)']^{-1}D(\iota)'H. \tag{12}$$

The value function under limited information simply replaces ζ with the best forecast $\bar{\zeta}$ of the state vector given past information on signals.

6.2 Full information

Consider first the full information environment in which states are observed and the model is known. The form of the value function is the same as that of Tallarini

(2000) and Barillas et al. (2007).[12] In a diffusion setting, a concern about robustness induces the consumer to consider distortions that append a drift $\mu_t dt$ to the Brownian increment and to impose a quadratic penalty to this distortion. This leads to a minimization problem whose indirect value function yields the T^1 operator of Hansen and Sargent (2007):

Problem 6.1.

$$0 = \min_{\mu} -\delta[\lambda(\iota) \cdot \zeta(\iota) + \kappa(\iota)] + \zeta(\iota)' D(\iota)' H + \mu' G(\iota)' H + \zeta(\iota)' A(\iota)' \lambda(\iota) \\ + \mu' B(\iota)' \lambda(\iota) + \frac{\theta_1}{2} \mu' \mu.$$

where we conjecture a value function of the form $\lambda(\iota) \cdot \zeta + \kappa(\iota) + c$*.*

Here θ_1 is a positive penalty parameter that characterizes the decision maker's fear that model ι is misspecified. We impose the same θ_1 for both models. See Hansen et al. (2006) and Anderson et al. (2003) for more general treatments and see appendix A for how we propose to calibrate θ_1. The minimizing drift distortion μ is

$$\mu^*(\iota) = -\frac{1}{\theta_1} \big[G(\iota)' H + B(\iota)' \lambda(\iota) \big] \tag{13}$$

which is independent of the state vector $\zeta(\iota)$. As a result,

$$\kappa(\iota) = -\frac{1}{2\theta_1 \delta} |G(\iota)' H + B(\iota)' \lambda(\iota)|^2 \tag{14}$$

Equating coefficients on $\zeta(\iota)$ in (6.1) implies that equation (11) continues to hold. Thus, $\lambda(\iota)$ remains the same as in the model without robustness and is given by (12).

Proposition 6.2. *The value function shares the same* $\lambda(\iota)$ *with the expected utility model [formula ((11))] and* $\kappa(\iota)$ *is given by* (14)*. The associated worst case distribution for the Brownian increment is normal with covariance matrix* Idt *and drift* $\mu^*(\iota) dt$ *given by* (13).

Under full information, the likelihood of the worst-case model relative to that of

[12] While Tallarini adopts an interpretation in terms of enhanced risk aversion, we interpret a risk-sensitivity adjustment as expressing an consumer's concern about model misspecification. See Barillas et al. (2007) for the relationship between these interpretations.

the benchmark model is a martingale $\{M_t^f(\iota)\}$ with local evolution

$$d\log M_t^f(\iota) = \mu^*(\iota)'dW_t - \frac{1}{2}|\mu^*(\iota)|^2 dt.$$

The stochastic discount factor (relative to the benchmark model) includes contributions both from the consumption dynamics and from the martingale, so that

$$d\log S_t^f = d\log M_t^f(\iota) - \delta dt - dc_t.$$

The local risk price is once again the negative of the risk exposure of the stochastic discount factor. With robustness, the risk price vector under full information is augmented by an uncertainty price:

$$\begin{array}{ccc} G(\iota)H & + & \frac{1}{\theta_1}[G(\iota)'H + B(\iota)'\lambda(\iota)] \\ \uparrow & & \uparrow \\ \text{risk} & & \text{uncertainty.} \end{array}$$

Neither the risk price nor the uncertainty price is state dependent or time dependent. We have completed the first row of Table 2.

6.3 Unknown states

Now suppose that the model (the value of ι) is known but the state $\zeta_t(\iota)$ is not. We seek a martingale $M_t^i(\iota)$ to use in (10) under this information structure. Following Hansen and Sargent (2007), we introduce a positive penalty parameter θ_2 and construct a robust estimate of the hidden state $\zeta_t(\iota)$ by solving:

Problem 6.3.

$$\min_{\phi \int \phi(\zeta)\psi(\zeta|\hat{\zeta},\Sigma)d\zeta=1} \int [\lambda(\iota)\cdot\zeta + \kappa(\iota) + \theta_2\phi(\zeta)\log\phi(\zeta)]\,\psi(\zeta|\hat{\zeta},\Sigma)d\zeta$$
$$= \min_{\tilde{\zeta}} \ \lambda(\iota)\cdot\tilde{\zeta} + \kappa(\iota) + \frac{\theta_2}{2}[\tilde{\zeta} - \bar{\zeta}(\iota)]'[\Sigma(\iota)]^{-1}[\tilde{\zeta} - \bar{\zeta}(\iota)]$$

where $\psi(\zeta|\bar{\zeta},\Sigma)$ is the normal density with mean $\bar{\zeta}$ and covariance matrix Σ, $\bar{\zeta}(\iota)$ is the estimate of state and Σ the covariance matrix under the benchmark ι model.

In the first line of Problem 6.3, ϕ is a density (relative to a normal) that distorts the density for the hidden state and θ_2 is a positive penalty parameter that penalizes ϕ's with large values of relative entropy (the expected value of $\phi \log \phi$). The second line of Problem 6.3 exploits the fact that the worst-case density is necessarily normal with a mean distortion $\tilde{z}$ to the state. This structure make it straightforward to compute the integral and as a result simplifies the minimization problem. In particular, the worst-case state estimate $\tilde{\zeta}(\iota)$ solves

$$0 = \lambda(\iota) + \frac{1}{\theta_2}[\Sigma(\iota)]^{-1}[\tilde{\zeta}(\iota) - \bar{\zeta}(\iota)].$$

Proposition 6.4. *The robust value function is*

$$U[\iota, \bar{\zeta}(\iota), \Sigma(\iota)] = \lambda(\iota) \cdot \bar{\zeta}(\iota) + \kappa(\iota) - \frac{1}{2\theta_2}\lambda(\iota)'\Sigma(\iota)\lambda(\iota) \tag{15}$$

with the same $\lambda(\iota)$ as in the expected utility model [formula (11)*] and the same $\kappa(\iota)$ as in the robust planner's problem with full information [formula* (14)*]. The worst-case state estimate is*

$$\tilde{\zeta} = \bar{\zeta} - \frac{1}{\theta_2}\Sigma(\iota)\lambda(\iota).$$

The indirect value function on the right side of (15) defines an instance of the T^2 operator of Hansen and Sargent (2007). Under the distorted evolution, dy_t has drift

$$\tilde{\xi}_t(\iota)dt = D(\iota)\tilde{\zeta}_t(\iota)dt + G(\iota)\mu^*(\iota)dt,$$

while under the benchmark evolution it has drift

$$\bar{\xi}_t(\iota)dt = D(\iota)\bar{\zeta}_t dt.$$

The corresponding likelihood ratio for our limited information setup is a martingale $M_t^i(\iota)$ that evolves as

$$d\log M_t^i(\iota) = [\tilde{\xi}_t(\iota) - \bar{\xi}_t(\iota)]'[\bar{G}(\iota)']^{-1}d\bar{W}_t(\iota) - \frac{1}{2}|\bar{G}(\iota)^{-1}[\tilde{\xi}_t(\iota) - \bar{\xi}_t(\iota)]|^2 dt,$$

and therefore the stochastic discount factor evolves as

$$d\log S_t^i = d\log M_t^i(\iota) - \delta dt - dc_t.$$

There are now two contributions to the uncertainty price, the one in the last column of the first row of table 2 coming from the potential misspecification of the state dynamics as reflected in the drift distortion to the Brownian motion and the other in the second row of table 2 coming from the filtering problem as reflected in a distortion in the estimated mean of hidden state vector:

$$\begin{array}{ccccc} \bar{G}(\iota)'H & + & \frac{1}{\theta_1}[\bar{G}(\iota)]^{-1}G(\iota)[G(\iota)'H + B(\iota)'\lambda(\iota)] & + & \frac{1}{\theta_2}[\bar{G}(\iota)]^{-1}D(\iota)\Sigma_t(\iota)\lambda(\iota) \\ \uparrow & & \uparrow & & \uparrow \\ \text{risk} & & \text{model uncertainty} & & \text{estimation uncertainty.} \end{array} \tag{16}$$

The state estimation adds time dependence to the uncertainty prices through the evolution of the covariance matrix $\Sigma_t(\iota)$ governed by (6), but the observed history of signals is inconsequential. We have completed the second row of Table 2.

6.4 Model unknown

Finally, we obtain a martingale M_t^u to use in pricing formula (10) that reflects a robust adjustment for an unknown model. We do this by twisting the model probability $\bar{\iota}_t$ by solving:

Problem 6.5.

$$\begin{aligned} \min_{0\le\tilde{\iota}\le 1} \quad & \tilde{\iota}U[1,\bar{\zeta}(1),\Sigma(1)] + (1-\tilde{\iota})U[0,\bar{\zeta}(0),\Sigma(0)] \\ & + \theta_2\tilde{\iota}[\log\tilde{\iota} - \log\bar{\iota}] + \theta_2(1-\tilde{\iota})[\log(1-\tilde{\iota}) - \log(1-\bar{\iota})] \end{aligned}$$

Proposition 6.6. *The indirect value function for this problem becomes our robust value function*[13]

$$-\theta_2 \log\left[\bar{\iota}\exp\left(-\frac{1}{\theta_2}U[1,\bar{\zeta}(1),\Sigma(1)]\right) + (1-\bar{\iota})\exp\left(-\frac{1}{\theta_2}U[0,\bar{\zeta}(0),\Sigma(0)]\right)\right].$$

The worst-case model probabilities satisfy:

$$\begin{aligned} (1-\tilde{\iota}) \quad &\propto \quad (1-\bar{\iota})\exp\left(-\frac{U[0,\bar{\zeta}(0),\Sigma(0)]}{\theta_2}\right) \\ \tilde{\iota} \quad &\propto \quad \bar{\iota}\exp\left(-\frac{U[1,\bar{\zeta}(1),\Sigma(1)]}{\theta_2}\right). \end{aligned}$$

[13]This is evidently another application of the T^2 operator of Hansen and Sargent (2007).

Under the distorted probabilities, the signal increment dy_t has a drift

$$\tilde{\kappa}_t dt = [\tilde{\iota}_t \tilde{\bar{\xi}}_t(1) + (1 - \tilde{\iota}_t)\tilde{\bar{\xi}}_t(0)]dt,$$

and under the benchmark probabilities this drift is

$$\bar{\kappa}_t dt = [\bar{\iota}_t \bar{\xi}_t(1) + (1 - \bar{\iota}_t)\bar{\xi}_t(0)]dt.$$

The associated martingale constructed from the relative likelihoods has evolution

$$d \log M_t^u = (\tilde{\kappa}_t - \bar{\kappa}_t)'(\bar{G}')^{-1} d\bar{W}_t - \frac{1}{2}|\bar{G}^{-1}(\tilde{\kappa}_t - \bar{\kappa}_t)|^2 dt$$

and the stochastic discount factor is

$$d \log S_t = d \log M_t^u - \delta dt - dc_t$$

The resulting risk price vector equals the exposure of $d \log S_t$ to $d\bar{W}_t$ and is

$$\begin{aligned}
&\bar{G}'H + \tilde{\iota}\bar{G}^{-1}\left[\frac{1}{\theta_1}G(1)G(1)'H + \frac{1}{\theta_1}G(1)B(1)'\lambda(1) + \frac{1}{\theta_2}D(1)\Sigma(1)\lambda(1)\right] \\
&\quad + (1 - \tilde{\iota})\bar{G}^{-1}\left[\frac{1}{\theta_1}G(0)G(0)'H + \frac{1}{\theta_1}G(0)B(0)'\lambda(0) + \frac{1}{\theta_2}D(0)\Sigma(0)\lambda(0)\right] \\
&\quad + (\tilde{\iota} - \bar{\iota})\bar{G}^{-1}\left[D(1)\bar{\zeta}(1) - D(0)'\bar{\zeta}(0)\right] \qquad (17)
\end{aligned}$$

This local price includes the risk price as well as the adjustments for model misspecification and robust state estimation given in (16) where the latter two contributions are weighted by the distorted probabilities $\tilde{\iota}$. It also includes an additional term that measures the difference in model probabilities before and after the distortion. The resulting overall risk prices now inherit signal dependence from the model and state probabilities and time dependence from the state covariance matrices $\Sigma(\iota)$. In summary, by completing the last column of table 2, we have characterized three sources of uncertainty prices.

7 Illustrating the mechanism

To highlight the forces that govern the component contributions of model uncertainty to risk prices in formula (17), we create a long-run risk model along the lines of Bansal and Yaron (2004) and Hansen et al. (2008a). Our models share the form

$$\begin{aligned} d\zeta_{1t} &= a(\iota)\zeta_{1t}(\iota) + \sigma_1(\iota)dW_{1t} \\ d\zeta_{2t} &= 0 \\ dy_t &= \zeta_{1t} + \zeta_{2t} + \sigma_2(\iota)dW_{2t} \end{aligned} \tag{18}$$

where $\zeta_{1t}(\iota), \zeta_{2t}(\iota)$ are scalars and W_{1t}, W_{2t} are scalar components of the vector Brownian motion W_t and where $\zeta_{20}(\iota) = \mu_y(\iota)$ is the unconditional mean of consumption growth for model ι. We use the following discrete time approximation to the state space system (5):

$$\begin{aligned} \zeta_{t+\tau}(\iota) - \zeta_t(\iota) &= \tau A(\iota)\zeta_t(\iota) + \sqrt{\tau}B(\iota)w_{t+\tau} \\ y_{t+\tau} - y_t &= \tau D(\iota)\zeta_t(\iota) + \sqrt{\tau}G(\iota)w_{t+\tau} \end{aligned} \tag{19}$$

where now $w_{t+\tau} \sim \mathcal{N}(0, I)$ is an iid shock. We set $\tau = 1$.

A small negative $a(\iota)$ coupled with a small $\sigma_1(\iota)$ captures long-run risks in consumption growth. Bansal and Yaron (2004) justify such a specification with the argument that it fits consumption growth approximately as well as, and is therefore difficult to distinguish from, an iid consumption growth model, which we know to fit the aggregate per capital U.S. consumption data well. We respect this argument by forming two models with the same values of the signal noise $\sigma_2(\iota)$ but that with differing values of $\sigma_1(\iota), \rho(\iota) = a(\iota)+1$, and $\mu_y(\iota) = \zeta_{20}(\iota)$, give *identical* values of the likelihood. We impose $\rho(1) = .99$ to capture a long run risk model, while the equally good fitting $\iota = 0$ model has $\rho = .36$.[14] Thus, we have constructed our two models

[14] The sample for real consumption of services and durables runs over the period 1948II-2008III. To fit model $\iota = 1$, we fixed $\rho = .99$ and estimated $\sigma_1 = .0004257, \sigma_2 = .0048177, \mu_y = .004545$. Fixing σ_2 equal to .0048177, we then found a values of $\rho = .36$ and associated values $\sigma_1 =, 0020455, \mu_y = .00478258$ that give virtually the same value of the likelihood. In this way, we construct two good fitting models that are difficult to distinguish, with model $\iota = 1$ being the long-run risk model and model ι much more closely approximating an iid growth model. Freezing the value of σ_2 at the above value, the maximum likelihood estimates are $\rho = .8179, \sigma_1 = .00131659, \mu_y = .00474011$. The

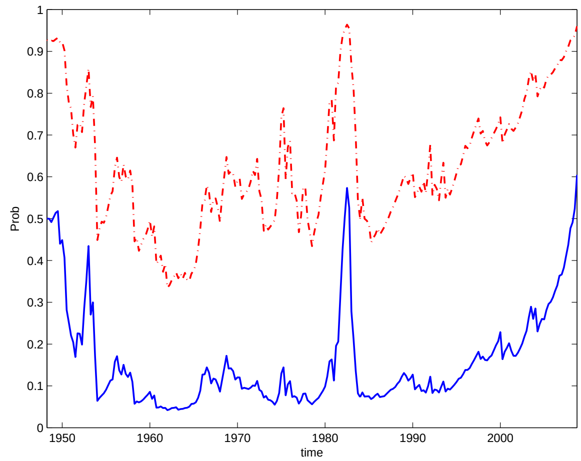

Figure 1: Bayesian model probability $\bar{\iota}_t$ (solid line) and worst-case model probability $\tilde{\iota}_t$ (dashed line).

so that they are indistinguishable statistically over our sample. This is our way of making precise the Bansal and Yaron (2004) observation that long-run risk and iid consumption growth models are difficult to distinguish empirically.

In appendix A we describe how we first calibrated θ_1 to drive the average detection error probability over the two ι models with observed states to be .4 and then with θ_1 thereby fixed set θ_2 to get a detection error probability of .2 for the signal distribution of the mixture model. We regard these values of detection error probabilities as being associated with moderate amounts of model uncertainty.[15] For these values of θ_1, θ_2,[16] figure 1 plots values of the Bayesian model mixing probability $\bar{\iota}$ along with the worst-

data for consumption comes from the St. Louis Fed data set (FRED). They are taken from their latest vintage (11/25/2008) with the following identifiers PCNDGC96_20081125 (real consumption on nondurable goods), PCESVC96_20081125 (real consumption on services). The population series is from the BLS, Series ID: LNS10000000. This is civilian noninstitutional population 16 years and over in thousands. The raw data are monthly. We averaged it to compute a quarterly series.

[15]We initiate the Bayesian probability $\bar{\iota}_0 = .5$ and set the covariance matrices $\Sigma_0(\iota)$ over hidden states at values that approximate what would prevail for a Bayesian who had previously observed a sample of the length 242 that we have in our actual sample period. In particular, we calibrated the initial state covariance matrices for both models as follows. First, we set preliminary 'uninformative' values that we took to be the variance of the unconditional stationary distribution of $\zeta_{2t}(\iota)$ and a value for the variance of $\zeta_2(\iota)$ of $.01^2$, which is orders of magnitude larger than the maximum likelihood estimates of μ_y for our entire sample. We set a preliminary state covariance between $\zeta_{1t}(\iota)$ and $\zeta_2(\iota)$ equal to zero. We put these preliminary values into the Kalman filter, ran it for a sample length of 242, and took the terminal covariance matrix as our starting value for the covariance matrix of the hidden state for model ι.

[16]The calibrated values are $\theta_1^{-1} = 7, \theta_2^{-1} = 1$.

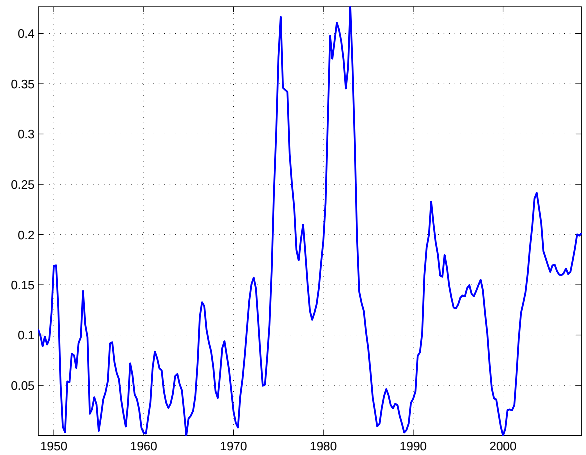

Figure 2: Contribution of model uncertainty to risk price.

case probability $\tilde{\iota}$. As described in the previous paragraph, we have constructed our two models so that with our setting of the initial model probability $\bar{\iota}_0$ at .5, the terminal value of $\bar{\iota}_t$ is also approximates .5. The interesting thing about figure 1 is to watch how the worst-case $\tilde{\iota}_t$ twists toward the long-run risk $\iota = 1$ model. This probability twisting contributes the countercyclical movements to the uncertainty contributions to risk prices from expression (17) that we plot in figure 2.[17]

7.1 Explanation for countercyclical uncertainty premia

Our representative consumer attaches positive probabilities to a model with statistically subtle persistence in consumption growth, namely, the long-run risk model of Bansal and Yaron (2004), and also to another model asserting close to iid consumption growth rates.[18] The asymmetrical response of model uncertainty premia to consumption growth shocks comes from (1) how consumer's concern about possible misspecification of the probabilities that he attaches to models causes him to calculate worst case probabilities that depend on value functions, and (2) how the value functions for the two models respond to shocks in ways that bring them closer together after positive consumption growth shocks and push them farther apart after negative shocks. The long-run risk model with very persistent consumption growth

[17]The figure plots all components of (17) except the ordinary risk price $\bar{G}H'$.

[18]Appendix B reports a sensitivity analysis aimed to add insight about the source of countercyclical risk prices.

confronts the consumer with a long-lived shock to consumption growth. That affects the set of possible model misspecifications that he worries about. The representative consumer's concerns about these misspecifications are reflected in a mean distortion to the long-run risk shock and therefore a more negative constant term in the formula for the continuation value. The resulting difference in constant terms in the value functions for the models with and without long-run consumption risk sets the stage for an asymmetric response of uncertainty premia to consumption growth shocks. Consecutive periods of higher than average consumption growth raise the probability that the consumer attaches to the model with persistent consumption growth relative to that of the approximately iid consumption growth $\iota = 0$ model. Although the long-run risk model has a more negative constant term, when a string of higher than average consumption growths occur, persistence of consumption growth under this model means that consumption growth can be expected to remain higher than average for many future periods. This pushes the continuation values associated with the two models closer together than they are when consumption growth rates have recently been lower than average. Via exponential twisting formulas, continuation values determine the worst-case probabilities that the representative consumer attaches to the models. That the continuation values for the two models become farther apart after a string of negative consumption growth shocks implies that our cautious consumer slants probability more towards the pessimistic long-run risk model when recent observations of consumption growth have been lower than average than when these observed growth rates have been higher than average. The intertemporal behavior of robustness-induced probability slanting accounts for how learning in the presence of uncertainty about models induces time variation in uncertainty premia.

7.2 Effects of learning under rational expectations

It is interesting to contrast the kind of pessimism coming from robustness in section 7 with the kind featured in Cogley and Sargent (2008) that is induced by a pessimistic prior joined with ordinary Bayesian learning. Figure 3 shows the contributions to risk prices $\gamma \bar{G}' H + \bar{G}^{-1} \left[\nu_t - D(\iota)\zeta_t \right]$ given in expression (8) when we assume that the true model used to price risks under full information is model $\iota = 0$ with parameters set at values estimated at the end of our sample. Notice how the learning contribution to the risk price fluctuates between positive and negative values. These fluctuations

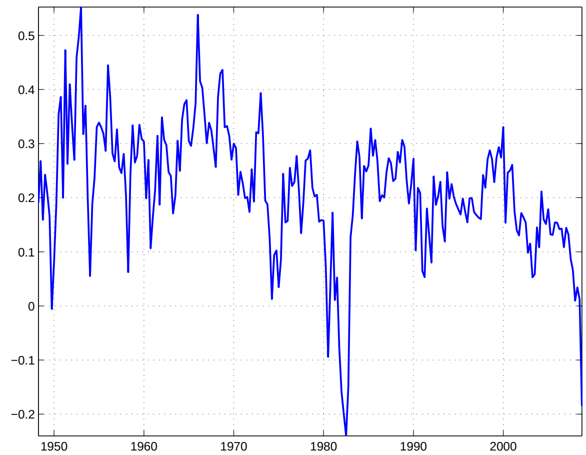

Figure 3: Contribution of learning to risk price.

that can be interpreted in terms of alternating spells of Bayesian-learning-induced optimism and pessimism relative to what we have assumed are the true hidden state variables with the true model.[19] The alternating signs of these effects of Bayesian learning on risk prices contrast with the unidirectional pessimism associated with model uncertainty.

8 Concluding remarks

The contributions of model uncertainty to risk prices combine (1) the same constant forward-looking contribution $\mu^*(\iota) = -\theta_1^{-1}\left[G(\iota)'H + B(\iota)'\lambda(\iota)\right]$ that was featured in earlier work without learning by Hansen et al. (1999) and Anderson et al. (2003), (2) additional smoothly decreasing in time components $-\theta_2^{-1}\Sigma(\iota)\lambda(\iota)$ that come from learning about parameter values within models, and (3) the potentially volatile time varying contribution highlighted in section 7.1 that is caused by the consumer's robust learning about the probability distribution over models.

[19] Suppose that the state vector processes $\{\zeta_t(\iota)\}$ are stationary and ergodic and the associated stationary distributions are used as the prior for the limited information structures. In this case, learning is about perpetually moving targets. In long samples, the entries of $\{\zeta_t(\iota) - \bar{\zeta}_t(\iota)\}$ will change signs so that on average they agree. In contrast if one an entry of $\zeta_t(\iota)$ is truly invariant but unknown *a priori*, then a systematic bias can emerge in a sample trajectory analogous to the one depicted in figure 3 even as the impact of the prior decays over time. For finite t's, the expectation of $\bar{\zeta}_t(\iota)$ conditioned on the invariant parameter will be biased as is standard in Bayesian analysis. This bias disappears only when we average across such trajectories in accordance to the prior over the invariant parameter.

Our mechanism for producing time varying risk premia differs from other approaches. For instance, Campbell and Cochrane (1999) induce secular movements in risk premia that are backward looking because a social externality depends on current and past average consumption. To generate variation in risk premia, Bansal and Yaron (2004) assume stochastic volatility in consumption.[20]

Our analysis features the effects of robust learning on local prices of exposure to uncertainty. Studying the consequences of robust learning and model selection for multi-period uncertainty prices is a natural next step. Multi-period valuation requires the compounding of local prices, and when the prices are time-varying this compounding can have nontrivial consequences.

Our analysis also imposed a unitary elasticity of substitution in order to obtain convenient formulas for prices. While a unitary elasticity of substitution simplifies our calculations, it implies that the ratio of consumption to wealth is constant. Although consumption claims have no obvious counterpart in financial data, it remains interesting to relax the unitary elasticity of substitution because of its potential importance in the valuation of durable claims.

While our example economy is highly stylized, we can imagine a variety of environments in which learning about low frequency phenomena is especially challenging when consumers are not fully confident about their probability assessments. Hansen et al. (2008a) show that while long-run risk components have important quantitative impacts on low frequency implications of stochastic discount factors and cash flows, it is statistically challenging to measure those components. Belief fragility emanating from model uncertainty promises to be a potent source of fluctuations in the prices of long-lived assets.

A Detection error probabilities

By adapting procedures developed by Hansen et al. (2002) and Anderson et al. (2003) in ways described by Hansen et al. (2008b), we can use simulations to approximate a detection error probability. Repeatedly simulate $\{y_{t+1} - y_t\}_{t=1}^{T}$ under the approximating model. Evaluate the likelihood functions the likelihood functions L_T^a and L_T^w of the approximating model and worst case model for a given (θ_1, θ_2). Compute

[20] Our interest in learning and time series variation in the uncertainty premium differentiates us from Weitzman (2005) and Jobert et al. (2006), who focus on long run averages.

the fraction of simulations for which $\frac{L_t^w}{L_T^a} > 1$ and call it r_a. This approximates the probability that the likelihood ratio says that the worst-case model generated the data when the approximating model actually generated the data. Do a symmetrical calculation to compute the fraction of simulations for which $\frac{L_T^a}{L_T^w} > 1$ (call it r_w), where the simulations are generated under the worst case model. As in Hansen et al. (2002) and Anderson et al. (2003), define the overall detection error probability to be

$$p(\theta_1, \theta_2) = \frac{1}{2}(r_a + r_w). \tag{20}$$

Because in this paper we use what Hansen et al. (2008b) call Game I, we use the following sequential procedure to calibrate θ_1 first, then θ_2. First, we pretend that $\zeta_t(\iota)$ is observable for $\iota = 0, 1$ and calibrate θ_1 by calculating detection error probabilities for a system with an observed state vector using the approach of Hansen et al. (2002) and Hansen and Sargent (2008, ch. 9). Then having pinned down θ_1, we use formula (20) to calibrate θ_2. This procedure takes the point of view that θ_1 measures how difficult it would be to distinguish one model of the partially hidden state from another if we were able to observe the hidden state, while θ_2 measures how difficult it is to distinguish alternative models of the hidden state. The probability $p(\theta_1, \theta_2)$ measures both sources of model uncertainty.

We proceeded as follows. (1) Conditional on model ι and the model ι state $\zeta_t(\iota)$ being observed, we computed the detection error probability as a function of θ_1 for models $\iota = 0, 1$. (2) Using a prior probability of $\pi = .5$, we averaged the two curves described in point (1) and plotted the average against θ_1. We calibrated θ_1 to yield an average detection error probability of .4 and used this value of θ_1 in the next step. (3) With θ_1 locked at the value just set, we then calculated and plotted the detection error for the mixture model against θ_2. To generate data under the approximating mixture model, we sampled sequentially from the conditional density of signals under the mixture model, building up the Bayesian probabilities $\bar{\iota}_t$ sequentially along a sample path. Similarly, to generate data under the worst case mixture model, we sampled sequentially from the conditional density for the worst-case signal distribution, building up the worst-case model probabilities $\tilde{\iota}_t$ sequentially. We set θ_2 to fix the overall detection error equal to .2.

B Sensitivity analysis

This appendix spotlights the force that produces countercyclical uncertainty contributions to risk prices by introducing a perturbation to our model that attenuates that force. The persistent countercyclical uncertainty contributions to risk prices in figure 3 come from a setting in which the representative consumer entertains two models that are difficult to distinguish. We study how uncertainty contributions to risk prices change when we expand the consumer's universe of models to include ones that fit the data even better than the two models in section 7. In particular, we now endow the model with seven models having the same value of σ_2 but now with values of $\rho = .36, .52, .67, .82, .89, .95, .99$, with values of σ_1, μ_y being concentrated out via likelihood function maximization. In terms of the likelihood function for the whole sample, the values at the end values .36, .99 are the poorest fitting ones and the $\rho = .82$ one is the best fitting. We start the representative consumer with a uniform prior over the seven models and set $\theta_1^{-1} = 4, \theta_2^{-1} = 3$ to obtain the uncertainty contribution to risk prices reported in figure 4. We report Bayesian model probabilities and their worst-case counterparts in figure 5.

Countercyclical risk premia still emerge, but they are moderated relative to those in figure 3 in the text. The reason is to be found in how the presence of models that fit better eventually pushes down the Bayesian model probability on long-run risk model. Pushing that probability down far enough diminishes its influence on uncertainty contributions to risk price even in the face of the tendency to twist model probabilities toward the long-run risk model. Even after twisting, the worst-case probabilities on that model are much smaller than they were in figure 1.

We find the comparison among competing models that have disperse implications for risk prices as featured in our paper to be interesting. While adding our models with ρ's between .36 and .99 that give higher values of the likelihood diminishes the variation in contributions to uncertainty that we have computed, the impact on the rational expectations counterpart can be even more dramatic because those intermediate models fit the data in ways that imply substantially smaller risk prices. For the recursive utility model with full information, the magnitude of the risk price is $\gamma|[B(\iota)'\lambda(\iota) + G(\iota)'H]|$ where γ is a measure of risk aversion. For the model with the largest likelihood, $|[B(\iota)'\lambda(\iota) + G(\iota)'H]| = 0.008622$, while this magnitude is 0.039627 for the $\rho = .99$ model. Thus, a value of γ more than four times larger is required at the maximum likelihood estimate for the magnitude of the risk price to remain the same as for the model with $\rho = .99$. To the extent that $\rho = .99$ is statistically implausible, a rational expectations econometrician either rejects the model or finds that consumers are highly risk averse.

References

Ai, Hengjie. 2006. Incomplete Information and Equity Premium in Production Economy. Unpublished.

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.

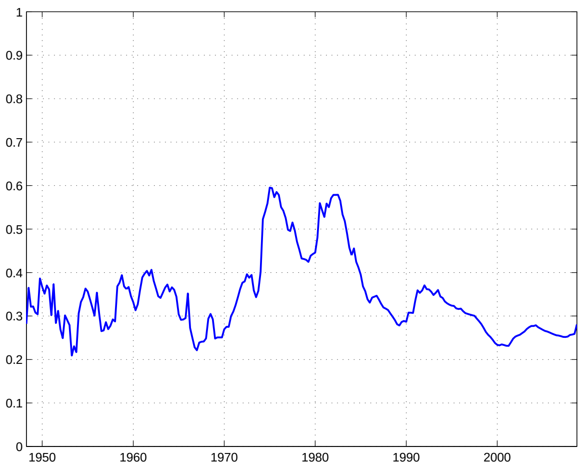

Figure 4: Contribution of model uncertainty to risk price with seven models.

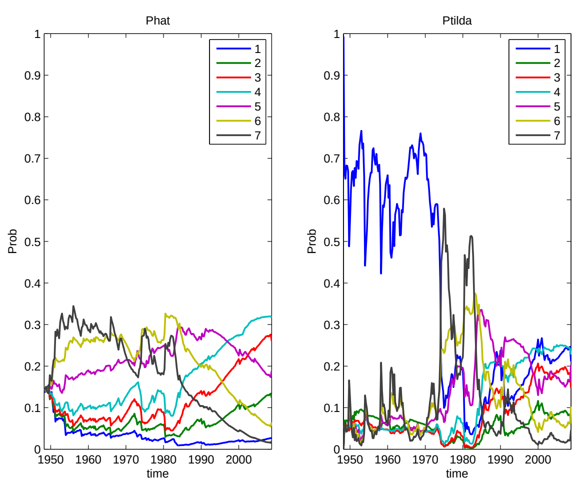

Figure 5: Bayesian and worst case model probabilities with seven models.

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

Barillas, Francisco, Lars Peter Hansen, and Thomas J. Sargent. 2007. Volatility or Doubts? Unpublished paper, University of Chicago and New York University.

Bossaerts, Peter. 2002. *The Paradox of Asset Pricing*. Princeton, NJ: Princeton University Press.

———. 2004. Filtering returns for unspecified biases in priors when testing asset pricing theory. *Review of Economic Studies* 71:6386.

Breeden, Douglas T. 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics* 7 (3):265 – 296.

Brennan, M. J. and Y. Xia. 2001. Stock Price Volatility and Equity Premium. *Journal of Monetary Economics* 47:249–283.

Cagetti, Marco, Lars Peter Hansen, Thomas J. Sargent, and Noah Williams. 2002. Robustness and Pricing with Uncertain Growth. *Review of Financial Studies* 15 (2):363–404.

Campbell, John Y. and John H. Cochrane. 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy* 107 (2):205–251.

Cerreia, Simone, Fabio Maccheroni, Massimo Marinacci, and Luigi Montrucchio. 2008. Uncertainty Averse Preferences. Working paper, Collegio Carlo Alberto.

Cogley, Timothy and Thomas J. Sargent. 2008. The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression? *Journal of Monetary Economics* 55:454–478.

Cogley, Timothy, Riccardo Colacito, Lars Hansen, and Thomas Sargent. 2005. Robustness and U.S. Monetary Policy Experimentation. Unpublished.

Croce, Mariano M., Martin Lettau, and Sydney C. Ludvigson. 2006. Investor Information, Long-Run Risk, and the Duration of Risky Cash Flows. Unpublished.

David, A. 1997. Fluctuating Confidence in Stock Markets. *Journal of Financial and Quantitative Analysis* 32:457–462.

Detemple, J. 1986. Asset Pricing in a Production Economy with Incomplete Information. *Journal of Finance* 41:383–390.

Duffie, Darrell. 2001. *Dynamic Asset Pricing Theory, Third Edition*. Princeton, New Jersey: Princeton University Press.

Gilboa, I. and D. Schmeidler. 1989. Maxmin Expected Utility with Non-Unique Prior. *Journal of Mathematical Economics* 18:141–153.

Hansen, L. P. and S. F. Richard. 1987. The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models. *Econometrica* 55:587–614.

Hansen, L. P. and T. J. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91:60–66.

Hansen, L. P., T. J. Sargent, G. A. Turmuhambetova, and N. Williams. 2006. Robust Control, Min-Max Expected Utility, and Model Misspecification. *Journal of Economic Theory* 128:45–90.

Hansen, Lars Peter. 2007. Beliefs, Doubts and Learning: Valuing Macroeconomic Risk. *American Economic Review* 97 (2):1–30.

———. 2008. Modeling the Long Run: Valuation in Dynamic Stochastic Economies. Presented as the Fisher-Schultz Lecture to the Econometric Societey.

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

Hansen, Lars Peter and Thomas J. Sargent. 2007. Recursive robust estimation and control without commitment. *Journal of Economic Theory* 136 (1):1–27.

———. 2008. *Robustness*. Princeton, New Jersey: Princeton University Press.

Hansen, Lars Peter and Jose Scheinkman. 2009. Long-term Risk: An Operator Approach. *Econometrica* 77 (1):177–234.

Hansen, Lars Peter, Thomas Sargent, and Thomas Tallarini. 1999. Robust Permanent Income and Pricing. *Review of Economic Studies* 66:873–907.

Hansen, Lars Peter, Thomas J. Sargent, and Neng E. Wang. 2002. Robust Permanent Income and Pricing with Filtering. *Macroeconomic Dynamics* 6:40–84.

Hansen, Lars Peter, John Heaton, and Nan Li. 2008a. Consumption Strikes Back: Measuring Long Run Risk. *Journal of Political Economy* 116 (2):260–302.

Hansen, Lars Peter, Ricardo Mayer, and Thomas J. Sargent. 2008b. Robust Hidden Markov LQG Problems. University of Chicago and New York University.

Harrsion, J. Michael and David M. Kreps. 1978. Speculative Investor Behavior in a Market with Heterogeneous Expectations. *Quarterly Journal of Economics* 92 (2):323–336.

Jobert, A., A. Platania, and L. C. G. Rogers. 2006. A Bayesian Solution to the Equity Premium Puzzle. Statistical Laboratory, University of Cambridge.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini. 2006a. Ambiguity Aversion, Robustness, and the Variational Representation of Preferences. *Econometrica* 74 (6):1447–1498.

———. 2006b. Dynamic Variational Preferences. *Journal of Economic Theory* 128:4–44.

Scheinkman, Jose A. and Wei Xiong. 2003. Overconfidence and Speculative Bubbles. *Journal of Political Economy* 111 (6):1183–1220.

Strzalecki, Tomasz. 2008. Axiomatic Foundations of Multiplier Preferences. Northwestern University Department of Economics.

Tallarini, Thomas D. 2000. Risk-sensitive real business cycles. *Journal of Monetary Economics* 45 (3):507–532.

Veronesi, P. 2000. How Does Information Quality Affect Stock Market Returns. *Journal of Finance* 55:807–837.

Weitzman, M. 2005. A Unified Theory of Equity Puzzles. Harvard University.