Available online at www.sciencedirect.com
ScienceDirect
Journal of Economic Theory 144 (2009) 2388–2418
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Doubts or variability?

Francisco Barillas [a], Lars Peter Hansen [b], Thomas J. Sargent [a,c,*]

[a] *New York University, United States*
[b] *University of Chicago, United States*
[c] *Hoover Institution, United States*

Received 9 August 2007; final version received 22 September 2008; accepted 26 November 2008

Available online 20 January 2009

Abstract

Reinterpreting most of the market price of risk as a price of model uncertainty eradicates a link between asset prices and measures of the welfare costs of aggregate fluctuations that was proposed by Hansen, Sargent, and Tallarini [17], Tallarini [30], Alvarez and Jermann [1]. Prices of model uncertainty contain information about the benefits of removing model uncertainty, not the consumption fluctuations that Lucas [22,23] studied. A max–min expected utility theory lets us reinterpret Tallarini's risk-aversion parameter as measuring a representative consumer's doubts about the model specification. We use model detection instead of risk-aversion experiments to calibrate that parameter. Plausible values of detection error probabilities give prices of model uncertainty that approach the Hansen and Jagannathan [11] bounds. Fixed detection error probabilities give rise to virtually identical asset prices as well as virtually identical costs of model uncertainty for Tallarini's two models of consumption growth.

JEL classification: C12; C52; E13; E30

Keywords: Risk aversion; Model misspecification; Robustness; Market price of risk; Equity premium puzzle; Risk-free rate puzzle; Detection error probability; Costs of model uncertainty

* Corresponding author at: Department of Economics, New York University, 19 W. Fourth Street, New York, NY, United States.

E-mail addresses: fvb201@nyu.edu (F. Barillas), l-hansen@uchicago.edu (L.P. Hansen), ts43@nyu.edu (T.J. Sargent).

0022-0531/$ – see front matter
doi:10.1016/j.jet.2008.11.014

No one has found risk aversion parameters of 50 or 100 in the diversification of individual portfolios, in the level of insurance deductibles, in the wage premiums associated with occupations with high earnings risk, or in the revenues raised by state-operated lotteries. It would be good to have the equity premium resolved, but I think we need to look beyond high estimates of risk aversion to do it.

Robert Lucas Jr., January 10, 2003

1. Introduction

In terms of their effects on asset prices and real quantities, can plausible concerns about robustness to model misspecification substitute for the implausibly large risk aversion parameters that bother Lucas in the above epigraph?[1] To answer this question, we reinterpret an elegant graph of Tallarini [30] by transforming Tallarini's CRRA risk-aversion parameter γ into a parameter that measures a set of probability models for consumption growth that are difficult to distinguish and over which a representative consumer seeks a robust valuation. To restrict γ, we use detection error probabilities that measure the proximity of probability distributions, as advocated by Anderson et al. [2] and Hansen and Sargent [16, Ch. 9], and we recast Tallarini's key diagram in terms of model detection error probabilities. A connection between model detection probabilities and a price of model uncertainty transcends specific approximating models. That price compensates the representative consumer for bearing model uncertainty, not risk.[2] We show that modest amounts of model uncertainty can substitute for large amounts of risk aversion in terms of choices and effects on asset prices.

Reinterpreting risk prices as model uncertainty prices makes them uninformative about the benefits of reducing aggregate fluctuations as defined by Lucas [22,23] and implies that those costs were mismeasured by Tallarini [30] and Alvarez and Jermann [1], who used connections between risk prices and costs of fluctuations that had been set forth by Hansen et al. [17]. To elaborate on this observation, we fashion a mental experiment about the welfare benefits from removing model uncertainty, an experiment that differs conceptually from Lucas's, but about which prices of model uncertainty are informative.

Section 2 reviews Hansen and Jagannathan's characterization of the equity premium and risk free rate puzzles that emerge with time separable CRRA preferences. Section 3 describes the stochastic setting and preferences that express aversion to model uncertainty. Sections 4 and 5 describe how Tallarini [30] used a preference of Kreps and Porteus [21] to find values of a risk-aversion parameters γ, one for a random walk model of log consumption, another for a trend

[1] Hansen et al. [17] describe a locus of (β, γ) pairs that are observationally equivalent for consumption and investment in linear-quadratic production economies, but that nevertheless imply different prices for risky assets. This finding is the basis of what Lucas [23, p. 7] calls Tallarini's [30] finding of "an astonishing separation of quantity and asset price determination...." Although this paper studies only pure endowment economies, the analytical observational equivalence result of Hansen et al. [17] and the approximate version of that result in Tallarini [30] make us confident that the theoretical values of quantities that will emerge from production economies will not be affected by alterations in the risk-sensitivity parameter γ that we use to measure concerns about model misspecification.

[2] See Anderson et al. [2].

stationary model, that can explain the risk-free rate puzzle of Weil [31]. But those values of γ are so high that they provoked Lucas's skeptical remark. Section 6 defines a concern about robustness to alternative models of log consumption growth that are constructed using martingale perturbations that Hansen and Sargent [13,15] and Hansen et al. [10] used to represent alternative specifications that are statistically near an approximating model. We then reinterpret Tallarini's utility recursion in terms of some max–min expected utility formulations in which the minimization operator expresses an agent's doubts about his stochastic specification. We describe senses in which risk aversion and model uncertainty aversion are and are not observationally equivalent. Section 7 reinterprets Tallarini's findings in terms of model uncertainty aversion. We use detection error probabilities to justify selecting different context-specific values of γ for the two approximating models of log consumption growth used by Tallarini, then modify Tallarini's key figure by recasting it in terms of detection probabilities. The figure reveals a link between the market price of model uncertainty and the detection error probability that transcends differences in the stochastic specification of the representative consumer's approximating model for log consumption growth, an outcome that could be anticipated from the tight relationship between the market price of model uncertainty and a large deviation bound on detection error probabilities derived by Anderson et al. [2]. Section 8 measures the benefits from a hypothetical experiment that removes model uncertainty, explains how this experiment differs from the mental experiment that underlies calculations of the benefits of reducing aggregate fluctuations by Lucas [22,23], Tallarini [30], and Alvarez and Jermann [1], and tells how the benefits of eliminating model uncertainty are reflected in the market price of model uncertainty. Section 9 discusses whether and how someone can learn not to fear model misspecification. Section 10 concludes.

Our analysis highlights how random shocks confront consumers with model ambiguity by obscuring differences among statistical models. Here random shocks discomfort consumers and investors in ways that Lucas [22,23], Tallarini [30], and Alvarez and Jermann [1] ignored.

2. The equity premium and risk-free rate puzzles

Along with Tallarini [30], we begin with a characterization of the risk-free rate and equity premium puzzles by Hansen and Jagannathan [11]. The random variable $m_{t+1,t}$ is said to be a stochastic discount factor if it confirms the following equation for the time t price p_t of a one-period payoff y_{t+1}:

$$p_t = E_t(m_{t+1,t} y_{t+1}),$$

where E_t denotes the mathematical expectation conditioned on date t information. For time-separable CRRA preferences with discount factor β, $m_{t+1,t}$ is simply the marginal rate of substitution:

$$m_{t+1,t} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \tag{1}$$

where C_t is consumption and γ is the coefficient of relative risk aversion. The reciprocal of the gross one-period risk-free rate is

$$\frac{1}{R_t^f} = E_t[m_{t+1,t}] = E_t\left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \right]. \tag{2}$$

Let ξ_{t+1} be the one-period excess return on a security or portfolio of securities. Using the definition of a conditional covariance and a Cauchy–Schwarz inequality, Hansen and Jagannathan [11] deduce the following bound:

Table 1
Sample moments from quarterly U.S. data 1948:1–2006:4, r^e is the real quarterly return on the value-weighted NYSE portfolio and r^f is the real quarterly return on the three month Treasury bill. Returns are measured in percent per quarter.

Return	Mean	Std. dev.
r^e	2.27	7.68
r^f	0.32	0.61
$r^e - r^f$	1.95	7.67
Market price of risk: 0.25		

$$\frac{|E_t[\xi_{t+1}]|}{\sigma_t(\xi_{t+1})} \leqslant \frac{\sigma_t(m_{t,t+1})}{E_t[m_{t,t+1}]}. \tag{3}$$

The left-hand side of (3) is the Sharpe ratio. The maximum Sharpe ratio is commonly called the market price of risk. It is the slope of the (conditional) mean-standard deviation frontier and is the increase in the expected rate of return needed to compensate an investor for bearing a unit increase in the standard deviation of return along the efficient frontier. The Sharpe ratio is bounded by the right-hand side of relation (3). With complete markets the bound is attained.

A counterpart to this inequality uses unconditional expectations and results in Hansen and Jagannathan's statement of the equity premium puzzle.[3] To reconcile formula (1) with measures of the market price of risk extracted from data on asset returns and prices only (like those in Table 1) requires a value of γ so high that it elicits doubts like those expressed by Lucas [23] in the epigraph starting this paper.[4,5]

But another failure isolated by the $\times$'s in Fig. 1 motivated Tallarini [30]. The figure plots an unconditional version of the Hansen and Jagannathan bound (the parabola) as well as the $\times$'s, which are pairs of unconditional mean $E(m)$ and the unconditional standard deviation $\sigma(m)$ implied by Eqs. (1) and (2) for different values of γ.[6] The figure addresses whether values of γ can be found for which the associated $(E(m), \sigma(m))$ pairs are inside the unconditional version of Hansen and Jagannathan bounds. The line of $\times$'s shows that high values of γ deliver high market prices of risk but also push the reciprocal of the risk-free rate down and therefore away from the Hansen and Jagannathan bounds. This is the risk-free rate puzzle of Weil [31].

In Section 5 we shall explain the loci of circles and crosses in Fig. 1. These loci depict how, by adopting a recursive preference specification, Tallarini [30] found values of γ that pushed $(E(m), \sigma(m))$ pairs inside the Hansen and Jagannathan bounds. That achievement registered as a mixed success because the values of γ that work are so high that, when interpreted as measures of risk-aversion, they provoked Lucas's skeptical remark.

[3] Conditioning information is brought in through the back door by scaling payoffs by variables in the conditioning information set and using an expanded collection of payoffs with prices that are one on average in place of gross returns.

[4] The "market price of risk" reported in Table 1 ignores conditioning information, but it remains a valid lower bound on the ratio of volatility of the intertemporal marginal rate of substitution relative to its mean.

[5] Precursors to Hansen and Jagannathan [11] are contained in Shiller [28] and a comment there by Hansen. Shiller deduced an inequality from the marginal distributions of consumption and returns while Hansen and Jagannathan [11] use the marginal distribution for the stochastic discount factor and the joint distribution for returns. Hansen and Jagannathan [11] thus featured maximal Sharpe ratios in their volatility bounds.

[6] For CRRA time-separable preferences, formulas for $E(m)$ and $\sigma(m)/E(m)$ are, first, for the random walk model, $E[m] = \beta \exp[\gamma(-\mu + \frac{\sigma_\varepsilon^2 \gamma}{2})]$ and $\frac{\sigma(m)}{E[m]} = \{\exp[\sigma_\varepsilon^2 \gamma^2] - 1\}^{1/2}$ and, second, for the trend stationary model $E[m] = \beta \exp[\gamma(-\mu + \frac{\sigma_\varepsilon^2 \gamma}{2}(1 + \frac{1-\rho}{1+\rho}))]$ and $\frac{\sigma(m)}{E[m]} = \{\exp[\sigma_\varepsilon^2 \gamma^2 (1 + \frac{1-\rho}{1+\rho})] - 1\}^{1/2}$.

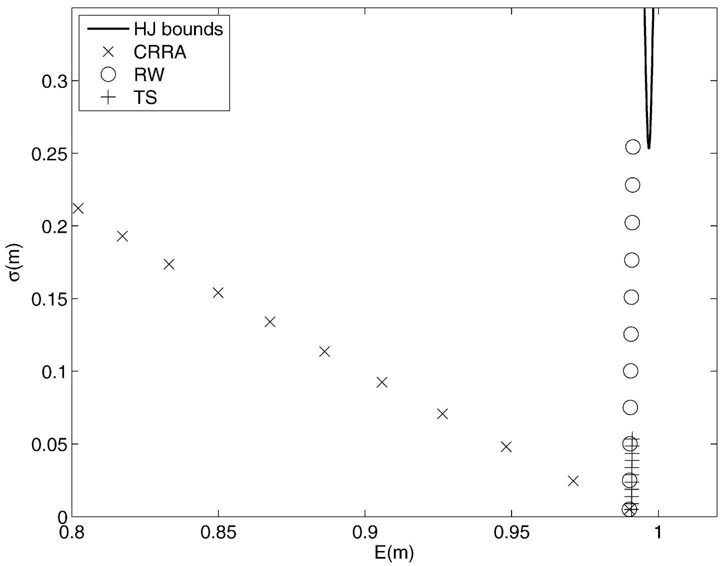


Fig. 1. *Solid line*: Hansen–Jagannathan volatility bound for quarterly returns on the value-weighted NYSE and Treasury Bill, 1948–2006. *Circles*: Mean and standard deviation for intertemporal marginal rate of substitution generated by Epstein–Zin preferences with random walk consumption. *Pluses*: Mean and standard deviation for stochastic discount factor generated by Epstein–Zin preferences with trend stationary consumption. *Crosses*: Mean and standard deviation for intertemporal marginal rate of substitution for CRRA time separable preferences. The coefficient of relative risk aversion, γ takes on the values 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 and the discount factor $\beta = 0.995$.

3. The choice setting

To prepare the way for Tallarini's findings and our reinterpretation of them, it is convenient to introduce the following objects in terms of which alternative decision theories are cast.

3.1. Shocks and consumption plans

We let $c_t = \log C_t$, x_0 be an initial state vector, $\varepsilon^t = [\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_1]$, and $\{\varepsilon_{t+1},\ t \geqslant 0\}$ be a sequence of random shocks with conditional densities $\pi_{t+1}(\cdot|\varepsilon^t, x_0)$ and an implied joint distribution $\Pi_\infty(\cdot|x_0)$ over the entire sequence. Let $\mathcal{C}$ be a set of consumption plans C^∞ whose time t elements C_t are measurable functions of (ε^t, x_0). Soon we shall consider a restricted class of consumption plans in $\mathcal{C}$ that have the following recursive representation:

$$\begin{aligned} x_{t+1} &= Ax_t + B\varepsilon_{t+1}, \\ c_t &= Hx_t \end{aligned} \tag{4}$$

where x_t is an $n \times 1$ state vector, ε_{t+1} is an $m \times 1$ shock, and the eigenvalues of A are bounded in modulus by $\frac{1}{\sqrt{\beta}}$. Representation (4) implies that the time t element of a consumption plan can be expressed as the following function of x_0 and the history of shocks:

$$c_t = H\big(B\varepsilon_t + AB\varepsilon_{t-1} + \cdots + A^{t-1}B\varepsilon_1\big) + HA^t x_0. \tag{5}$$

We let $\mathcal{C}(A, B, H; x_0)$ denote the set of consumption plans with representation (4)–(5).

Table 2
Estimates from quarterly U.S. data 1948:2–2006:4. Standard errors in parentheses.

Parameter	Random walk	Trend stationary
μ	0.00495	0.00418
	(0.0003)	(0.0003)
σ_ε	0.0050	0.0050
	(0.0002)	(0.0002)
ρ	–	0.980
		(0.010)
ζ	–	−4.48
		(0.08)

In this paper, we use one of the following two consumption plans that Tallarini finds fit post WWII U.S. per capita consumption well:

1. Geometric random walk:

$$c_t = c_0 + t\mu + \sigma_\varepsilon(\varepsilon_t + \varepsilon_{t-1} + \cdots + \varepsilon_1), \quad t \geqslant 1, \tag{6}$$

where

$$\varepsilon_t \sim \pi_t\big(\cdot|\epsilon^{t-1}, x_0\big) = \pi(\cdot) \sim \mathcal{N}(0,1).$$

2. Geometric trend stationary[7]:

$$c_t = \rho^t c_0 + \mu t + \big(1-\rho^t\big)\zeta + \sigma_\varepsilon\big(\varepsilon_t + \rho\varepsilon_{t-1} + \cdots + \rho^{t-1}\varepsilon_1\big), \quad t \geqslant 1, \tag{7}$$

where

$$\varepsilon_t \sim \pi_t\big(\cdot|\epsilon^{t-1}, x_0\big) = \pi(\cdot) \sim \mathcal{N}(0,1).$$

3.2. Parameter estimates

We estimated both consumption processes using U.S. quarterly real consumption growth per capita from 1948:2–2006:4.[8] Maximum likelihood point estimates are summarized in Table 2. We shall use these point estimates as inputs into the calculations below.

3.3. Overview of agents I, II, III, and IV

The preferences of our four types of agent over consumption plans $C^\infty \in \mathcal{C}$ are defined in terms of the following sets of objects:

Type I agent (Kreps–Porteus–Epstein–Zin–Tallarini):

(i) a discount factor $\beta \in (0,1)$; (ii) an intertemporal elasticity of substitution IES equal to unity; (iii) a risk aversion parameter $\gamma \geqslant 1$; and (iv) a conditional density $\pi_{t+1}(\cdot|\varepsilon^t, x_0) = \pi(\cdot)$ for ε_{t+1} and an implied joint distribution $\Pi_\infty(\cdot|x_0)$.

[7] The recursive version of our trend stationary model is $c_t = \zeta + \mu t + z_t$, $z_t = \rho z_{t-1} + \varepsilon_t$.

[8] Consumption is measured as real personal consumption expenditures on nondurable goods and services and is deflated by its implicit chain price deflator. We use the same deflator to deflate asset prices. We use civilian noninstitutional population 16 years and older to get per capita series.

Type II agent (ambiguity averse Hansen and Sargent [9] multiplier preferences):

(i) a discount factor $\beta \in (0,1)$; (ii) an intertemporal elasticity of substitution IES equal to unity; (iii) a risk aversion parameter equal to 1; (iv) a conditional density $\pi_{t+1}(\cdot|\varepsilon^t, x_0)$ for ε_{t+1} and an implied joint distribution $\Pi_\infty(\cdot|x_0)$; and (v) a parameter θ that penalizes the entropy associated with a minimizing player's perturbation of Π_∞ relative to the iid, standard normal benchmark.

Type III agent (ambiguity averse Hansen and Sargent [9] constraint preferences):

(i) a discount factor $\beta \in (0,1)$; (ii) an intertemporal elasticity of substitution IES equal to unity; (iii) a risk aversion parameter equal to 1; (iv) a conditional density $\pi_{t+1}(\cdot|\varepsilon^t, x_0)$ for ε_{t+1} and an implied joint distribution $\Pi_\infty(\cdot|x_0)$; and (v) a parameter η that measures the discounted relative entropy of perturbations to $\Pi_\infty(\cdot|x_0)$ relative to an iid, standard normal benchmark allowable to a minimizing player.

Type IV agent (pessimistic *ex post* Bayesian):

(i) a discount factor $\beta \in (0,1)$; (ii) an IES $= 1$; (iii) a risk-aversion parameter of 1; and (iv) a unique pessimistic joint probability distribution $\hat{\Pi}_\infty(\cdot|x_0, \theta)$.

Our reinterpretation of Tallarini's quantitative findings as well as our mental experiment that measures the costs of model specification uncertainty both hinge on the following behavioral implications of these alternative preference specifications. Agents I and II are observationally equivalent in the strong sense that they have identical preferences over $\mathcal{C}$. Agents III and IV are observationally equivalent with I and II in the more restricted, but for us still very useful, sense that their valuations of risky assets coincide at an exogenous endowment process that we take to be the approximating model for the type II and type III representative agents.

4. A type I agent: Kreps–Porteus–Epstein–Zin–Tallarini

Our type I agent has preferences over $\mathcal{C}$ that are defined via the value function recursion

$$\log V_t = (1-\beta)c_t + \beta \log\big[E_t(V_{t+1})^{1-\gamma}\big]^{\frac{1}{1-\gamma}} \tag{8}$$

where $\gamma \geqslant 1$. This is the risk-sensitive recursion of Hansen and Sargent [12,15] that for a logarithmic period utility function Tallarini [30] interpreted to be a case of the recursive preference specification of Epstein and Zin [6,7] in which the intertemporal elasticity of substitution IES is fixed at unity and the atemporal coefficient of relative risk aversion is γ.

4.1. Formulas for continuation values

To represent asset prices, we first compute continuation values for the two alternative consumption processes. Define $U_t \equiv \log V_t/(1-\beta)$ and

$$\theta = \frac{-1}{(1-\beta)(1-\gamma)}. \tag{9}$$

Then

$$U_t = c_t - \beta\theta \log E_t\left[\exp\left(\frac{-U_{t+1}}{\theta}\right)\right]. \tag{10}$$

When $\gamma = 1$ (or $\theta = +\infty$), recursion (10) becomes the standard discounted expected utility recursion

$$U_t = c_t + \beta E_t U_{t+1}.$$

For consumption processes $C^\infty \in \mathcal{C}(A, B, H; x_0)$ associated with different specifications of (A, B, H) in (4), recursion (10) implies the following Bellman equation[9]:

$$U(x) = c - \beta\theta \log \int \exp\left[\frac{-U(Ax + B\varepsilon)}{\theta}\right]\pi(\varepsilon)\,d\varepsilon. \tag{11}$$

For the random walk specification, the solution of (10) is

$$U_t = \frac{\beta}{(1-\beta)^2}\left[\mu - \frac{\sigma_\varepsilon^2}{2\theta(1-\beta)}\right] + \frac{1}{1-\beta}c_t. \tag{12}$$

For the trend stationary model the solution of the value function recursion (10) is:

$$U_t = \frac{\beta\zeta(1-\rho)}{(1-\beta)(1-\beta\rho)} + \frac{\beta\mu}{(1-\beta)^2} - \frac{\sigma_\varepsilon^2\beta}{2\theta(1-\beta)(1-\beta\rho)^2} + \frac{\mu\beta(1-\rho)}{(1-\beta\rho)(1-\beta)}t + \frac{1}{1-\beta\rho}c_t. \tag{13}$$

4.2. Pricing implications

Arrow securities are defined relative to a measure used to integrate over states. We first use the Lebesgue measure. For a type I representative agent economy, the price of a one-period Arrow security is

$$\left(\beta\frac{C_t}{C_{t+1}(\varepsilon^*)}\right)\left(\frac{\exp[-U_{t+1}(\varepsilon^*)/\theta]}{\int \exp[-U_{t+1}(\varepsilon)/\theta]\,d\pi(\varepsilon)}\right)\pi(\varepsilon^*).$$

We abuse notation to avoid proliferation. The notation C_{t+1} is the random variable that denotes consumption at date $t+1$. Recognizing that the new information available at date $t+1$ is captured by the random vector ε_{t+1}, $C_{t+1}(\cdot)$ explicitly represents the dependence of C_{t+1} on ε_{t+1} and similarly for $U_{t+1}(\cdot)$.

Instead of using the Lebesgue measure to integrate over states, the stochastic discount factor uses the underlying conditional probability distribution. As a consequence, the stochastic discount factor is given by

$$m_{t+1,t} = \left(\beta\frac{C_t}{C_{t+1}}\right)\left(\frac{\exp(-U_{t+1}/\theta)}{E_t[\exp(-U_{t+1}/\theta)]}\right). \tag{14}$$

The change in the reference measure for integration leads to π being omitted in (14).

In conjunction with a solution for the value function, for example, (12) or (13), Eq. (14) shows how the standard stochastic discount factor $(\beta\frac{C_t}{C_{t+1}})$ associated with time separable logarithmic utility is altered by a potentially volatile function of the continuation value U_{t+1} when the risk aversion parameter $\gamma \equiv 1 + \frac{1}{(1-\beta)\theta} > 1$ (see Eq. (9)). For a type I agent, γ is a risk aversion parameter that differs from the reciprocal of the IES when $\gamma \neq 1$.

[9] The notation U_t denotes the continuation value realized at date t for a consumption plan. In Bellman equation (11), we use $U(\cdot)$ to denote the value function as a function of the Markov state.

5. A type I agent economy with high risk aversion attains HJ bound

For the random walk and trend stationary consumption processes, Tallarini computed the following formulas for $E(m)$ and $\sigma(m)$ for what we call a type I agent. For the random walk model that follows, the mean and volatility of m_{t+1} conditioned on date t information are constant. For the trend stationary models they depend on conditioning information, and we will work with their unconditional counterparts.

- Random walk model:

$$E[m] = \beta \exp\left[-\mu + \frac{\sigma_\varepsilon^2}{2}(2\gamma - 1)\right], \tag{15}$$

$$\frac{\sigma(m)}{E[m]} = \left\{\exp\left[\sigma_\varepsilon^2 \gamma^2\right] - 1\right\}^{\frac{1}{2}}. \tag{16}$$

- Trend stationary model:

$$E[m] = \beta \exp\left[-\mu + \frac{\sigma_\varepsilon^2}{2}\left(1 - \frac{2(1-\beta)(1-\gamma)}{1-\beta\rho} + \frac{1-\rho}{1+\rho}\right)\right], \tag{17}$$

$$\frac{\sigma(m)}{E[m]} = \left\{\exp\left[\sigma_\varepsilon^2\left(\left\{\frac{(1-\beta)(1-\gamma)}{1-\beta\rho} - 1\right\}^2 + \frac{1-\rho}{1+\rho}\right)\right] - 1\right\}^{\frac{1}{2}}. \tag{18}$$

Fig. 1 is our version of Tallarini's key figure. It follows Tallarini in using the above formulas to plot loci of $(E(m), \sigma(m))$ pairs as the risk-aversion parameter γ varies.[10] This figure chalks up a striking success for Tallarini compared to the corresponding risk-free-rate-puzzle laden $\times$'s in Fig. 1 for time separable CRRA preferences. Notice how for both specifications of the endowment process, increasing γ pushes the volatility of the stochastic discount factor upward toward the Hansen–Jagannathan bound while leaving $E(m)$ essentially unaffected, thus avoiding the risk-free rate puzzle of Weil [31].[11]

However, to approach the Hansen–Jagannathan bound Tallarini had to set the risk aversion parameter γ to very high values, 50 for the random walk model, about 250 for the trend stationary model. These high values provoked the skeptical remarks we have cited from Lucas [23].

6. Reinterpretations

We respond to Lucas's reluctance to use Tallarini's findings as a source of evidence about a representative consumer's attitude about random consumption fluctuations by reinterpreting γ as a parameter that expresses model specification doubts rather than risk aversion.

[10] As observed by Kocherlakota [20], for the random walk model, it is possible to generate the $(E(m), \sigma(m))$ pairs in Fig. 1 while sticking with the time separable CRRA model by changing β along with γ in the following way: $(\gamma, \beta) =$ (1, 0.9950), (5, 1.0147), (10, 1.0393), (15, 1.0637), (20, 1.0881), (25, 1.1124), (30, 1.1364), (35, 1.1602), (40, 1.1838), (45, 1.2070), (50, 1.2300).

[11] By comparing the formulas for $E(m)$ in footnote 6 with formula (15) for $E(m)$ for the random walk case, one sees how by locking the IES equal to 1, formula (15) arrests the force in the footnote 6 equation that pushes downward $E(m)$ as one increases γ via the term $\exp(-\gamma\mu)$. For power utility this is the dominant effect of γ on $E(m)$ when μ is much larger than σ_ε, as it is in the data.

6.1. Language for robustness: an 'approximating model'

To express doubts about model specification, we put multiple probability specifications on the table. To stay as close as possible to rational expectations, we work with a setting in which a representative agent has one fully specified model represented with particular A, B, H and Π_∞ in (4). In this paper, that model will be either the random walk model or the trend stationary model described above. We shall call this the 'approximating model' to acknowledge that the agent does not completely trust it. We express specification doubts in terms of alternative joint distributions Π that the agent contemplates assigning to the shocks ε^∞. We imagine that the agent surrounds his approximating model with a set of unspecified models that are statistically nearby (as measured by conditional relative entropy) and that he thinks might govern the data. Our type II and III agents want one value function that will somehow let them evaluate consumption plans under all of those nearby models. Before telling how they get those value functions, we first describe a mathematical formalism for representing the unspecified densities over ε^∞ that concern the agent as statistical perturbations of $\Pi_\infty(\varepsilon^\infty)$.

6.2. Using martingales to represent probability distortions

Let the representative consumer's information set be $\mathcal{X}_t$, which for us will be the history of log consumption growth rates up to date t. Random variables that are $\mathcal{X}_t$ measurable can be expressed as Borel measurable functions of x_0 and ε^t. Hansen and Sargent [13,15] use a nonnegative $\mathcal{X}_t$-measurable function G_t with $E(G_t|x_0) = 1$ to create a distorted probability measure that is absolutely continuous with respect to the probability measure over $\mathcal{X}_t$ generated by one of our two approximating models for log consumption growth.[12] Under the original probability measure the random variable G_t is a martingale with mean 1. We can use G_t as a Radon–Nikodym derivative (i.e., a likelihood ratio) to generate a distorted measure under which the expectation of a bounded $\mathcal{X}_t$-measurable random variable W_t is $\tilde{E}W_t \doteq EG_tW_t$. The *entropy* of the distortion at time t conditioned on date zero information is $E(G_t \log G_t|\mathcal{X}_0)$.

6.3. Recursive representations of distortions

We often factor a joint density F_{t+1} over an $\mathcal{X}_{t+1}$-measurable random vector as $F_{t+1} = f_{t+1}F_t$, where f_{t+1} is a one-step ahead density conditioned on $\mathcal{X}_t$. Following Hansen and Sargent [13], it is also useful to factor G_{t+1}. Form

$$g_{t+1} = \begin{cases} \frac{G_{t+1}}{G_t} & \text{if } G_t > 0, \\ 1 & \text{if } G_t = 0. \end{cases}$$

Then $G_{t+1} = g_{t+1}G_t$ and

$$G_t = G_0 \prod_{j=1}^{t} g_j. \tag{19}$$

The random variable G_0 is equal to unity. By construction, g_{t+1} has date t conditional expectation equal to unity. For a bounded random variable b_{t+1} that is $\mathcal{X}_{t+1}$-measurable, the distorted conditional expectation implied by the martingale $\{G_t\colon t \geqslant 0\}$ is

[12] See Hansen et al. [10] for a continuous time formulation.

$$\frac{E(G_{t+1}b_{t+1}|\mathcal{X}_t)}{E(G_{t+1}|\mathcal{X}_t)} = \frac{E(G_{t+1}b_{t+1}|\mathcal{X}_t)}{G_t} = E(g_{t+1}b_{t+1}|\mathcal{X}_t)$$

provided that $G_t > 0$. We extend this distorted conditional expectation to a more general collection of random variables by approximating unbounded random variables with a sequence of bounded ones. For each $t \geqslant 0$, construct the space $\mathcal{G}_{t+1}$ of all nonnegative, $\mathcal{X}_{t+1}$-measurable random variables g_{t+1} for which $E(g_{t+1}|\mathcal{X}_t) = 1$. We use g_{t+1} to represent distortions of the conditional probability distribution for $\mathcal{X}_{t+1}$ given $\mathcal{X}_t$.

6.4. A type II agent: ambiguity averse multiplier preferences

We represent ambiguity aversion with the multiplier preferences of Hansen and Sargent [9] and Hansen et al. [10].[13] These are defined in terms of a parameter θ that penalizes the discrepancy between perturbed models and the approximating model and that is linked via an application of the Lagrange multiplier theorem to a parameter η that occurs in what we shall call the "constraint preferences" of our type III representative agent.

A type II agent's *multiplier* preference ordering over $C^\infty \in \mathcal{C}$ is described by

$$\min_{\{g_{t+1}\}} \sum_{t=0}^{\infty} E\big\{\beta^t G_t\big[c_t + \beta\theta E\big(g_{t+1}\log g_{t+1}|\varepsilon^t, x_0\big)\big]\big|x_0\big\} \tag{20}$$

where

$$G_{t+1} = g_{t+1}G_t, \qquad E\big[g_{t+1}|\varepsilon^t, x_0\big] = 1, \quad g_{t+1} \geqslant 0,\ G_0 = 1. \tag{21}$$

In this paper, we restrict ourselves to studying subsets $\mathcal{C}(A, B, H; x_0)$ of $\mathcal{C}$ with typical element C^∞. For this set of consumption plans $C^\infty \in \mathcal{C}(A, B, H; x_0)$, a type II agent has a value function

$$W(x_0) = \min_{\{g_{t+1}\}} \sum_{t=0}^{\infty} E\big\{\beta^t G_t\big[c_t + \beta\theta E\big(g_{t+1}\log g_{t+1}|\varepsilon^t, x_0\big)\big]\big|x_0\big\} \tag{22}$$

where the minimization is subject to (4) and (21). The value function solves the following Bellman equation:

$$GW(x) = \min_{g(\varepsilon)\geqslant 0} G\bigg(c + \beta\int\big[g(\varepsilon)W(Ax + B\varepsilon) + \theta g(\varepsilon)\log g(\varepsilon)\big]\pi(\varepsilon)\,d\varepsilon\bigg). \tag{23}$$

Dividing by G gives

$$W(x) = c + \min_{g(\varepsilon)\geqslant 0}\bigg(\beta\int\big[g(\varepsilon)W(Ax + B\varepsilon) + \theta g(\varepsilon)\log g(\varepsilon)\big]\pi(\varepsilon)\,d\varepsilon\bigg)$$

[13] Maccheroni et al. [24,25] give an axiomatic foundation for variational preferences and describe how they express ambiguity aversion. Both multiplier and constraint preferences are special cases of variational preferences. Constraint preferences are particular instances of the multiple priors model of Gilboa and Schmeidler [8]. Strzalecki [29] and Cerreia et al. [3] give axiomatic defenses for multiplier preferences. Hansen and Sargent [15] show the link between the smooth ambiguity formulation of Klibanoff et al. [19] and multiplier preferences, while Cerreia et al. [3] show that the multiplier preferences used here are the only preferences that are both variational and smooth in the sense of Klibanoff et al. [19]. Specifically, smooth variational preferences necessarily use discrepancies between distributions measured with relative entropy.

where the minimization is subject to $\int g(\varepsilon)\pi(\varepsilon)\,d\varepsilon = 1$. Solving the minimum problem and substituting the minimizer into the above equation gives the risk-sensitive recursion of Hansen and Sargent [12,15]:

$$W(x) = c - \beta\theta \log \int \exp\left[\frac{-W(Ax + B\varepsilon)}{\theta}\right]\pi(\varepsilon)\,d\varepsilon. \tag{24}$$

The minimizing martingale increment is

$$\hat{g}_{t+1} = \left(\frac{\exp(-W(Ax_t + B\varepsilon_{t+1})/\theta)}{E_t[\exp(-W(Ax_t + B\varepsilon_{t+1})/\theta)]}\right). \tag{25}$$

6.5. *Types I and II are observationally equivalent*

Notice that Eqs. (11) and (24) imply that

$$W(x) \equiv U(x). \tag{26}$$

Therefore, agents I and II have identical preferences over elements of $C^\infty \in \mathcal{C}(A, B, H; x_0)$.[14] In this strong sense, they are observationally equivalent, but the interpretation of θ differs for type I and type II agents. For a type I agent, $\theta(\gamma) \equiv \frac{-1}{(1-\beta)(1-\gamma)}$ is a measure of risk aversion. For a type II agent, θ indicates his fear of model misspecification as measured by how much the minimizing agent gets penalized for raising entropy.

6.6. *A type III agent: ambiguity averse* constraint *preferences*

Hansen and Sargent [9,13] and Hansen et al. [10] describe *constraint preferences* that are directly related to the multiple priors model of Gilboa and Schmeidler [8]. Here a primitive object is a *set* of probability densities that we attribute to the representative type III consumer. We follow our earlier work by using ideas from robust control theory to construct this set of densities. In particular, we follow Hansen and Sargent [9,13] and restrain the discounted relative entropy of perturbations to the approximating model:

$$\beta E\left[\sum_{t=0}^{\infty} \beta^t G_t E\big(g_{t+1} \log g_{t+1} | \varepsilon^t, x_0\big)\Big| x_0\right] \leqslant \eta \tag{27}$$

where $\eta \geqslant 0$ measures the size of an entropy ball surrounding the distribution $\Pi_\infty(\varepsilon^\infty | x_0)$. Given a set of models within an entropy ball $\eta > 0$, constraint preferences over $C^\infty \in \mathcal{C}$ are ordered by

$$\min_{\{g_{t+1}\}} \sum_{t=0}^{\infty} E\big[\beta^t G_t c_t \big| x_0\big] \tag{28}$$

where the minimization is subject to $G_{t+1} = g_{t+1} G_t$ and $G_0 = 1$. If we restrict C^∞ to be in $\mathcal{C}(A, B, H; x_0)$, a type III agent has a value function

$$J(x_0) = \min_{\{g_{t+1}\}} \sum_{t=0}^{\infty} E\big[\beta^t G_t c_t \big| x_0\big] \tag{29}$$

[14] It can also be established that they have identical preferences for $C^\infty \in \mathcal{C}$.

where the minimization is subject to the discounted entropy constraint (27) and

$$\begin{aligned} &x_{t+1} = Ax_t + B\varepsilon_{t+1}, \\ &c_t = Hx_t, \quad x_0 \text{ given}, \\ &G_{t+1} = g_{t+1}G_t, \qquad E\big[g_{t+1}|\varepsilon^t, x_0\big] = 1, \quad g_{t+1} \geqslant 0,\ G_0 = 1. \end{aligned} \tag{30}$$

Hansen and Sargent [9], Hansen et al. [10], and Hansen and Sargent [16, Ch. 6] describe how constraint and multiplier preferences differ and also how θ and η can be chosen to align choices and valuations along equilibrium paths of the associated two-player zero-sum games. Briefly, they show how (1) *ex post* θ in the multiplier preferences can be viewed as the Lagrange multiplier on the time 0 discounted entropy constraint, and (2) the multiplier θ and continuation entropy can be chosen to align equilibrium outcomes that emerge from multiplier and constraint preferences.

6.7. *A type IV agent:* ex post *Bayesian*

A type IV agent is an ordinary expected utility agent with log preferences and a particular (distorted) joint distribution $\hat{\Pi}_\infty(\cdot|x_0)$ over C^∞:

$$\hat{E}_0 \sum_{t=0}^{\infty} \beta^t c_t. \tag{31}$$

The joint distribution $\hat{\Pi}_\infty(\cdot|x_0)$ is the one associated with the preferences of a type II agent and so depends on θ as well as on A, B, H when we restrict C^∞ to lie within $\mathcal{C}(A, B, H; x_0)$. The value function for a type IV agent equals the value function $J(x)$ for a type III agent.

6.8. *Types III and IV not observationally equivalent to I or II, but ...*

While agents I and II have identical preference orderings over $\mathcal{C}(A, B, H; x_0)$ (and more generally over $C^\infty \in \mathcal{C}$), they have different preference orderings than agents III and IV. Still, there is a more limited but for us very important sense in which agents of all four types look alike. It is true that for a *fixed* $\hat{\pi}(\varepsilon^\infty|A, B, H, \theta)$, the type IV pessimistic agent makes different choices over $\mathcal{C}(A, B, H; x_0)$ (i.e., plans C^∞ represented in terms of $(\tilde{A}, \tilde{B}, \tilde{H}) \neq (A, B, H)$) than does the type I or type II agent. However, for the *particular* A, B, H plan and θ used to derive the worst-case joint distribution $\hat{\Pi}(\varepsilon^\infty)$, the shadow prices of uncertain claims for a type IV agent match those for a type II agent.[15] This provides an interesting perspective on what are ordinarily interpreted as prices of "risk" in settings in which no one fears model uncertainty.

6.9. *Interpretation of stochastic discount factor*

The same representation of the stochastic discount factor

$$m_{t+1,t} = \left(\beta \frac{C_t}{C_{t+1}}\right)\hat{g}\big(\varepsilon(t+1)\big) \tag{32}$$

15 This link extends to decision problems in which one can apply the Minimax Theorem. See Hansen et al. [10] for more discussion.

prevails for all four types of representative consumer, but the interpretation of $\hat{g}$ varies across the four types. With a type I representative consumer in the style of Tallarini [30], the distortion $\hat{g}(\varepsilon_{t+1})$ is a contribution from the Kreps and Porteus [21] recursive utility specification that gets intermediated through continuation values; $\hat{g}(\varepsilon_{t+1})$ deviates from being identically unity when risk aversion $\gamma > 1$ exceeds the inverse of the IES. For the max–min expected utility type II and III representative agents, $\hat{g}(\varepsilon_{t+1})$ is the likelihood ratio that transforms the one-step conditional density $\pi(\epsilon_{t+1})$ under the approximating model to the worst-case density that these consumers use to evaluate risky streams. The fact that $\hat{g}(\varepsilon_{t+1})\pi(\varepsilon_{t+1})$ is the (unique) subjective conditional density for ε_{t+1} for a type IV *ex post* Bayesian representative agent means that as outside analysts we must introduce the likelihood ratio $\hat{g}(\varepsilon_{t+1})$ into the stochastic discount factor whenever we want to use the approximating model to price assets.

6.10. *Choice of data generating model*

The sets of probability models that surround the approximating model and that help define the preferences of the type II and III agents are in their heads. To make empirical statements, we have to posit a data generating model. The rational expectations hypothesis assumes that there is a unique distribution, i.e., that all subjective distributions equal a presumed objective one, so that after a rational expectations is formulated, the data generating mechanism is not an extra object to specify. But since we have put multiple probability models into the heads of our types II and III agents, and a unique pessimistic one into the head of our type IV agent, we have to make an explicit assumption about a unique data generating mechanism. We assume that the approximating model is the data generating mechanism, so the fears of model misspecification of our type II and III agents are, after all, only in their heads. Having taken this stance, we shall use the Radon–Nikodym derivative $\hat{g}$ to price the model uncertainty feared by our type II and III representative consumers.

6.11. *Value functions and discounted entropy*

In terms of the minimizing martingale increment we can express the value function recursion for a type II agent as

$$W(x) = c + \beta \int \big[\hat{g}(\varepsilon) W(Ax + B\varepsilon) + \theta \hat{g}(\varepsilon) \log \hat{g}(\varepsilon)\big]\pi(\varepsilon)\, d\varepsilon. \tag{33}$$

By solving (33), we can express $W(x)$ as the sum of two components, the first of which is the expected discounted value of C^∞ under the worst case model, while the second is θ times discounted entropy:

$$W(x) = J(x) + \theta N(x) \tag{34}$$

where

$$J(x) = c + \beta \int \big[\hat{g}(\varepsilon) J(Ax + B\varepsilon)\big]\pi(\varepsilon)\, d\varepsilon \tag{35}$$

and

$$N(x) = \beta \int \big[\hat{g}(\varepsilon) \log \hat{g}(\varepsilon) + \hat{g}(\varepsilon) N(Ax + B\varepsilon)\big]\pi(\varepsilon)\, d\varepsilon. \tag{36}$$

Here

$$J(x_t) = \hat{E}_t \sum_{j=0}^{\infty} \beta^j c_{t+j}$$

is the expected discounted log consumption under the worst case joint density for C^∞ and

$$G_t N(x_t) = G_t \beta E\left[\sum_{j=0}^{\infty} \beta^j \frac{G_{t+j}}{G_t} E\left[g_{t+j+1} \log g_{t+j+1} | \varepsilon^{t+j}, x_0\right] \Big| \varepsilon^t, x_0 \right]$$

is continuation entropy.

While $W(x)$ is the value function for a type II agent, evidently $J(x)$ is the value function for a type III agent as well as for a type IV agent. To evaluate $W(x)$ and $J(x)$ it is useful first to find the minimizing martingale increment $\hat{g}(\varepsilon)$ and then discounted entropy. We do that in the next two subsections.

6.12. Minimizing martingale increments

Using formula (25), we find that the minimizing martingale increment for the geometric random walk model is:

$$\hat{g}_{t+1} \propto \exp\left(\frac{-\sigma_\varepsilon \varepsilon_{t+1}}{(1-\beta)\theta}\right).$$

The implied distorted conditional density is

$$\hat{\pi}(\varepsilon_{t+1}) \propto \exp\left(\frac{-\varepsilon_{t+1}^2}{2}\right) \exp\left(\frac{-\sigma_\varepsilon \varepsilon_{t+1}}{(1-\beta)\theta}\right).$$

Completing the square gives

$$\hat{\pi}(\varepsilon) \sim \mathcal{N}\big(w(\theta), 1\big) \tag{37}$$

where

$$w(\theta) = \frac{-\sigma_\varepsilon}{(1-\beta)\theta}. \tag{38}$$

Pursuing an analogous calculation for the trend stationary model, we find that the worst case conditional density again has form (37), where now

$$w(\theta) = -\frac{\sigma_\varepsilon}{(1-\rho\beta)\theta}. \tag{39}$$

6.13. Discounted entropy

When the conditional densities for ε_{t+1} under the approximating and worst case models are $\pi \sim \mathcal{N}(0, 1)$ and $\hat{\pi} \sim \mathcal{N}(-w(\theta), 1)$, respectively, we can compute that conditional entropy is

$$E_t \hat{g}_{t+1} \log \hat{g}_{t+1} = \int \big(\log \hat{\pi}(\varepsilon) - \log \pi(\varepsilon)\big) \hat{\pi}(\varepsilon)\, d\varepsilon = \frac{1}{2} w'(\theta) w(\theta).$$

It then follows that discounted entropy becomes

$$\beta E\left[\sum_{t=0}^{\infty} \beta^t \hat{G}_t E\big(\hat{g}_{t+1} \log \hat{g}_{t+1} | \varepsilon^t, x_0\big) \Big| x_0 \right] = \eta = \frac{\beta}{2(1-\beta)} w'(\theta) w(\theta). \tag{40}$$

Formula (40) gives a mapping between θ and η that allows us to set these parameters to align multiplier and constraint preferences along an exogenous endowment process. We shall use this mapping to interpret θ below. In particular, after we introduce detection error probabilities in Section 7, we shall argue that it is more natural to fix η rather than θ when we make comparisons by altering the consumer's baseline approximating model from the random walk to the trend stationary model. For this purpose, it is useful to note that by using formulas (38), (39), and (40), we find that the following choices of θ's for the random walk and trend stationary models imply identical discounted entropies[16]:

$$\theta_{\mathrm{TS}} = \left(\frac{\sigma_\varepsilon^{\mathrm{TS}}}{\sigma_\varepsilon^{\mathrm{RW}}}\right)\frac{1-\beta}{1-\rho\beta}\theta_{\mathrm{RW}}. \tag{41}$$

6.14. *Value functions for random walk log consumption*

Using the formula for $w(\theta)$ from the random walk model (38) tells us that discounted entropy is

$$N(x) = \frac{\beta}{2(1-\beta)}\frac{\sigma_\varepsilon^2}{(1-\beta)^2\theta^2}. \tag{42}$$

For the random walk model, we can then compute the value function for a type II agent to be

$$W(x_t) = \frac{\beta}{(1-\beta)^2}\left[\mu - \frac{\sigma_\varepsilon^2}{2(1-\beta)\theta}\right] + \frac{1}{1-\beta}c_t \tag{43}$$

and for a type III agent to be

$$J(x_t) = \frac{\beta}{(1-\beta)^2}\left[\mu - \frac{\sigma_\varepsilon^2}{(1-\beta)\theta}\right] + \frac{1}{1-\beta}c_t, \tag{44}$$

so that $W(x_t) = J(x_t) + \theta N(x_t)$. To interpret $J(x_t)$ as the value function for a type III agent, we use formula (40) to align θ and η. We shall use these value functions to construct compensating variations in the initial condition for log consumption c_0 in an elimination-of-model-uncertainty experiment to be described in Section 8.

6.15. *Market prices of risk and model uncertainty*

Hansen et al. [17,18] note that the conditional standard deviation of the Radon–Nikodym derivative $\hat{g}(\varepsilon)$ is

$$\mathrm{MPU} = \mathrm{std}_t(\hat{g}) = \left[\exp\bigl(w(\theta)'w(\theta)\bigr) - 1\right]^{\frac{1}{2}} \approx \left|w(\theta)\right|. \tag{45}$$

By construction $E_t\hat{g} = 1$. We call $\mathrm{std}_t(g)$ the market price of model uncertainty (MPU). It can be verified that for the random-walk and trend-stationary models, $|w_{t+1}|$ given by the above formulas comprises the lion's share of what Tallarini [30] interpreted as the market price of risk given by formulas (16) and (18). This is because the first difference in the log of consumption has a small conditional coefficient of variation in our data (this observation is the heart of the equity premium puzzle). Thus, formula (45) is a good approximation to Tallarini's formulas (16) and (18). It follows from formula (40) that

[16] The ratio $\sigma_\varepsilon^{\mathrm{TS}}/\sigma_\varepsilon^{\mathrm{RW}} = 1$ at the parameter values in Table 2.

$$\text{MPU} = \sqrt{\frac{2\eta(1-\beta)}{\beta}}.$$

6.16. Interpretation of MPU

As the slope of the mean-standard deviation frontier, the market price of risk (MPR) tells the increase in the expected return needed to compensate an investor for accepting a unit increase in the standard deviation of the return along the efficient frontier. Our type II and III consumers' worst-case beliefs encode their concerns about model misspecification. We can measure the market price of model uncertainty (MPU) in terms of how a representative investor's worst case model distorts mean returns. When measured using the approximating model, the worst case model's distortion in mean rates of return amplifies objects that are usually interpreted as a market price of risk. In a continuous time limit, the MPU is the maximal expected rate of return distortion in the worst case model relative to the approximating model, per unit of standard deviation of return. For example, see Anderson et al. [2].

7. Reinterpreting Tallarini

Tallarini interprets γ as a parameter measuring aversion to atemporal gambles. The quote from Lucas [23] and the reasoning of Cochrane [4], who applied ideas of Pratt [27], tell why economists think that only small positive values of γ are plausible when it is interpreted as a risk-aversion parameter. The mental experiment of Pratt confronts a decision maker with choices between gambles with *known* probability distributions (i.e., the type of risks that the type I agent thinks he faces).

The observational equivalence between our types I and II agents means that we can just as well interpret γ as measuring the consumer's concern about model misspecification. But how should we think about plausible values of γ (or θ) when it is to be interpreted as encoding responses to gambles that involve *unknown* probability distributions? We answer this question by using detection error probabilities that tell how difficult it is to distinguish probability distributions on the basis of a fixed finite number of observations. These measures inspire us to argue that it is not appropriate to regard γ or θ as a parameter that remains fixed when we vary the stochastic process for consumption under the consumer's approximating model, e.g., the random walk or trend stationary model for log consumption. Instead, we shall see that it is more plausible to fix the size of the discounted entropy ball η as we think of moving across approximating models. This is because the detection error probabilities turn out to be functions of η that vary little across the trend stationary and random walk models. Thus, our mental experiment under model uncertainty leads us to use the same values of the discounted entropy constraint η or the implied detection error probabilities, but different values of γ, for different approximating models.

7.1. Calibrating γ using detection error probabilities

This section describes how to use Bayesian detection error probabilities to calibrate a plausible value for γ or θ when it is interpreted as a parameter measuring a representative consumer's concern about model misspecification.[17] The idea is that it is plausible for agents to be concerned

[17] Also see Anderson et al. [2] and Hansen and Sargent [16].

about models that are difficult to distinguish from one another with data sets of moderate size. We implement this idea by focusing on statistically distinguishing the approximating model (call it model A) from a worst case model associated with a particular θ (call it model B). Imagine that before seeing any data, the agent had assigned probability 0.5 to both the approximating model and the worst case model associated with θ. After seeing T observations, the representative consumer performs a likelihood ratio test for distinguishing model A from model B. If model A were correct, the likelihood ratio could be expected falsely to say that model B generated the data p_A percent of the time. Similarly, if model B were correct, the likelihood ratio could be expected falsely to say that model A generated the data p_B percent of the time. We weight p_A, p_B by the prior probabilities 0.5 to obtain what we call the detection error probability:

$$p(\theta^{-1}) = \frac{1}{2}(p_A + p_B). \tag{46}$$

The detection error probability $p(\theta^{-1})$ is a function of θ^{-1} because the worst-case model depends on θ. When $\gamma = 1$ (or $\theta^{-1} = 0$, see Eq. (9)), it is easy to see that $p(\theta^{-1}) = 0.5$ because then the approximating and worst-case models are identical. As we raise θ^{-1} above zero, $p(\theta^{-1})$ falls below 0.5.

We use introspection to instruct us about plausible values of $p(\theta^{-1})$ as a measure of concern about model misspecification. Thus, we think it is sensible for a decision maker to want to guard against possible misspecifications whose detection error probabilities are 0.2 or even less.

As a function of θ^{-1}, $p(\theta^{-1})$ differs for different specifications of the approximating model. In particular, it will change when we switch from a trend stationary to a random walk model of log consumption. When comparing outcomes across different approximating models, we advocate comparing outcomes for the *same* detection error probabilities $p(\theta^{-1})$ and adjusting the θ^{-1}'s appropriately across models. We shall do that for our version of Tallarini's model and will recast his Fig. 1 in terms of loci that record $(E(m), \sigma(m))$ pairs as we vary the detection error probability.

7.2. *Tallarini's figure again*

The left panel of Fig. 2 describes the detection probability $p(\theta^{-1})$ for the random walk (dashed line) and trend stationary (solid line) models. We simulated the approximating and worst-case models 100,000 times and followed the procedure described above to compute the detection error probabilities for a given θ^{-1}. The simulations were done for $T = 235$ periods, the sample size for quarterly consumption growth data over the period from 1948II–2006IV.

The left panel of Fig. 2 reveals that for the random walk and the trend stationary models, a given detection error probability $p(\theta^{-1})$ is associated with different values of θ^{-1}. Therefore, if we want to compute $E(m), \sigma(m)$ pairs for the same detection error probabilities, we have to use different values of θ^{-1} for our two models of log per capita consumption growth. We shall use Fig. 2 to find these different values of θ^{-1} associated with a given detection error probability, then redraw Tallarini's figure in terms of detection error probabilities.

The right panel of Fig. 2 plots the detection error probabilities against the values of discounted entropy η for each model for the random walk and trend stationary models. As functions of η, the detection error probabilities for the two models are the same.

Thus, to prepare a counterpart to Fig. 1, our updated version of Tallarini's graph, we invert the detection error probability functions $p(\theta^{-1})$ in the left panel of Fig. 2 to get θ^{-1} as a function of

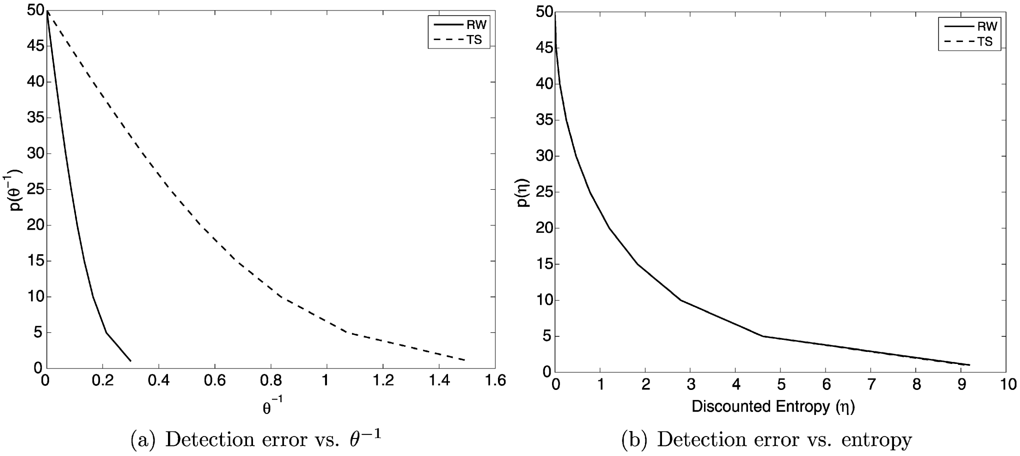


Fig. 2. Panel A: detection error probabilities versus θ^{-1} for the random walk and trend stationary models. Panel B: detection error probabilities versus discounted entropy η for the random walk and trend stationary models (the two curves coincide).

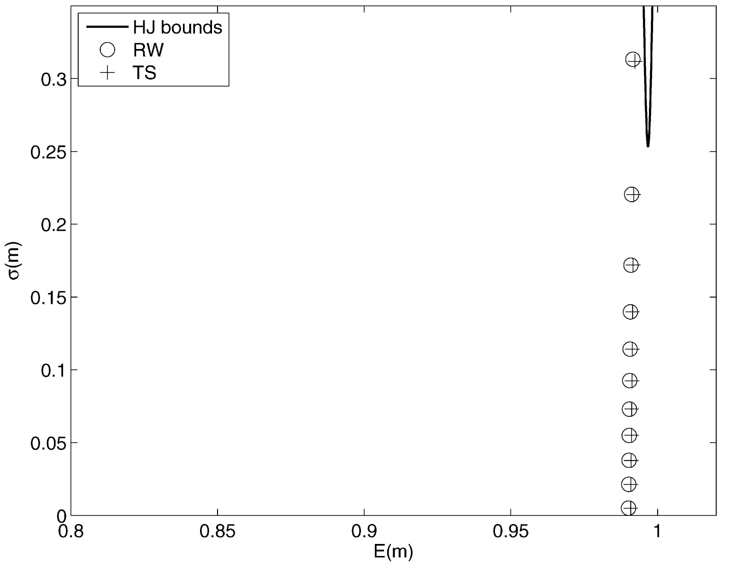


Fig. 3. Reciprocal of risk free rate, market price of risk pairs for the random walk (∘) and trend stationary (+) models for values of $p(\theta^{-1})$ of 50, 45, 40, 35, 30, 25, 20, 15, 10, 5 and 1 percent.

$p(\theta^{-1})$ for each model, then use this θ^{-1} either in formulas (15), (16) or in formulas (17), (18) to compute the $(E(m), \sigma(m))$ pairs to plot à la Tallarini. We present the results in Fig. 3.

We invite the reader to compare our Fig. 3 with Fig. 1. The calculations summarized in Fig. 1 taught Tallarini that with the random walk model for log consumption, the $(E(m), \sigma(m))$ pairs approach the Hansen and Jagannathan bound when γ is around 50, whereas under the trend stationary model we need γ to be 75 in order to approach the bound when $\beta = 0.995$. Fig. 3 simply repackages those results by using the detection error probabilities $p(\theta^{-1})$ reported in the left panel of Fig. 2 to trace out loci of $(E(m), \sigma(m))$ pairs as we vary the detection error probability.

Fig. 3 reveals the striking pattern that varying the detection error probabilities traces out nearly the same loci for the random walk and the trend stationary models of consumption. This outcome faithfully reflects a pattern that holds exactly for large deviation bounds on the detection error probabilities that were studied by Anderson et al. [2]. Their work established a tight link between those bounds and the market price of model uncertainty that transcends details of the stochastic specification for the representative consumer's approximating model.

In terms of the issue raised in the quote from Lucas [23], Fig. 3 reveals that regardless of the stochastic specification for consumption, what we regard as conservative detection error probabilities of between 0.15 and 0.2 take us half of the way toward the Hansen and Jagannathan bound.

To appreciate the significance of this finding, recall that Tallarini [30] showed how to explain both the equity premium and the risk-free rate by using Epstein–Zin–Weil preferences to separate a CRRA parameter γ from an IES parameter that he fixed at 1. To make things work, Tallarini needed very different levels of risk aversion depending on whether he used a random-walk with drift or a trend stationary model for log consumption. In Fig. 1, Tallarini needed to set $\gamma = 50$ for the random walk model and $\gamma = 75$ for the trend stationary model. For that figure, we follow Tallarini in setting $\beta = 0.995$, which implies an $E(m)$ whose inverse does not match the risk free rate in the economy very well: notice that in our Fig. 3, the circles and crosses lie a bit to the left of the Hansen–Jagannathan bound.[18]

Fig. 3 reveals that for the same detection error probability both models of consumption growth imply the same values of (what is ordinarily interpreted as) the market price of risk. We say "what is ordinarily interpreted as" in order to indicate that on our preferred interpretation, the contribution from $\hat{g}$, which accounts for most of it, should be interpreted as a "market price of model uncertainty." Fig. 3 alters our sense of how plausible a given setting of γ is when we see that one gets pretty close to the bound with a detection error probability of 5 percent. A representative consumer who sets a detection error probability that small does not seem to be as timid as one who sets a CRRA coefficient as high as 50 or 75.

8. Welfare gains from eliminating model uncertainty

Obstfeld [26], Dolmas [5], and Tallarini [30] studied the welfare costs of business cycle with Epstein–Zin preferences, while Hansen et al. [17] described links between asset prices and welfare costs of consumption fluctuations in settings that featured both risk-sensitivity and robustness. In this section, we revisit welfare calculations under our robustness interpretation instead of Tallarini's risk-sensitivity interpretation.

We have argued that the lion's share of what Tallarini [30] and Alvarez and Jermann [1] interpret as market prices of *risk* should instead be interpreted as market prices of *uncertainty*. This means that those uncertainty prices reveal the representative consumer's attitude about a very different mental experiment than the one that interested Lucas [22,23]. The question posed

[18] We can get inside the Hansen–Jagannathan bound by increasing the discount factor β. But, doing so requires even higher levels of the coefficient of risk aversion, especially for the trend stationary model. Adjusting the parameters in this way pushes the circles and pluses in our Fig. 3 to the right as we increase the discount factor. The level of detection error probability necessary to achieve a certain market price of model uncertainty is almost unaltered when we alter the discount factor.

by Lucas [22,23] was "how much consumption would the representative consumer be willing to sacrifice in order to avoid facing the risk associated with a *known* distribution of consumption fluctuations?" In the epigraph above, Lucas doubts that useful measures of the representative consumer's attitudes toward the type of macroeconomic risk that he had in mind can be recovered from asset market prices and returns by adopting the risk-sensitive interpretations of risk-premia in Hansen et al. [17], Tallarini [30], and Alvarez and Jermann [1].

In this section, we describe how market prices of uncertainty extracted from asset market data contain information about how much the representative consumer would be willing to pay to eliminate model uncertainty.

8.1. Comparison with risk-free certainty equivalent path under logarithmic random walk

In the spirit of Lucas [22], we follow Tallarini [30] by using as our point of comparison the certainty equivalent plan

$$c_{t+1} - c_t = \mu + \frac{1}{2}\sigma_\varepsilon^2. \tag{47}$$

We seek an adjustment to initial consumption, and therefore the scale of the entire process, that renders a representative consumer indifferent between the certainty equivalent plan and the original risky consumption plan. For the same initial conditions, the certainty equivalent path of consumption $\exp(c_{t+1})$ has the same mean as for the original plan $c_{t+1} - c_t = \mu + \sigma_\varepsilon \varepsilon_{t+1}$, but its conditional variance has been reduced to zero. We let c_0^J denote the level of initial log consumption in the certainty equivalent plan for a type J agent, where J is I, II, III or IV.

8.1.1. Type I agent

Recall formulas (12) and (43) for the value functions of type I and II representative agents facing a random walk process for log consumption, namely

$$U(x_0) = W(x_0) = \frac{\beta}{(1-\beta)^2}\left[\mu - \frac{\sigma_\varepsilon^2}{2(1-\beta)\theta}\right] + \frac{1}{1-\beta}c_0.$$

We seek a proportional decrease in the certainty equivalent trajectory (47) that leaves U equal to its value under the risky process. Let c_0^I denote the initialization of the certainty equivalent trajectory for a type I agent. Evidently, it satisfies the equation:

$$\frac{\beta}{(1-\beta)^2}\left(\mu + \frac{\sigma_\varepsilon^2}{2}\right) + \frac{1}{1-\beta}c_0^I = \frac{\beta}{(1-\beta)^2}\left[\mu - \frac{\sigma_\varepsilon^2}{2(1-\beta)\theta}\right] + \frac{1}{1-\beta}c_0.$$

The left side is the value under the certainty equivalent plan, while the right side is the value under the original risky plan starting from c_0. Solving for $c_0 - c_0^I$ gives

$$c_0 - c_0^I = \frac{\beta}{(1-\beta)}\left[\frac{\sigma_\varepsilon^2}{2} + \frac{\sigma_\varepsilon^2}{2(1-\beta)\theta}\right] \tag{48}$$

$$= \frac{\beta\sigma_\varepsilon^2}{2(1-\beta)}\left[1 + \frac{1}{(1-\beta)\theta}\right] \tag{49}$$

$$= \frac{\beta\sigma_\varepsilon^2\gamma}{2(1-\beta)}. \tag{50}$$

8.1.2. Type II agent

Because the value functions U and W for type I and II agents are identical, the compensating variation (50) renders both types indifferent between the original risky process and the certainty equivalent path. Thus $c_0^I = c_0^{II}$. But the reasons for indifference differ for the two types of agent. For our Kreps–Porteus type I agent, expression (50) makes risk aversion, as measured by γ, the reason that the consumer is willing to accept a lower initialization of the consumption path in order to eliminate volatility in the growth of the logarithm of consumption. However, for our type II agent with multiplier preferences indexed by θ, the reduction in initial consumption contains contributions from both risk-aversion and aversion to model uncertainty. The compensation $c_0 - c_0^{II}$ emerges from comparing what a type II agent regards as a trajectory that is both risky and model-uncertain with a trajectory that is both risk-free and model-certain. By itself, this comparison does not allow us to distinguish responses to risk and model uncertainty. To separate the parts contributed by risk and uncertainty, we construct a certainty equivalent for another type II agent, but one who does not fear model misspecification. This certainty equivalent starts from $c_0^{II}(r)$ instead of c_0^{II}.

Thus, consider a type II agent who does not fear model uncertainty, so that $\theta = +\infty$. We ask how much adjustment in the initial condition of a certainty equivalent path (47) a $\theta = +\infty$ type II consumer would require.[19] For $\theta = +\infty$, (50) asserts a compensating variation for the elimination of risk alone of

$$c_0 - c_0^{II}(r) = \frac{\beta\sigma_\varepsilon^2}{2(1-\beta)}. \tag{51}$$

For the random walk model, (51) corresponds to the compensation formula that Lucas [22] computed for a consumer with time separable logarithmic preferences, i.e., the special case of the preferences used by Tallarini [30] for a consumer whose coefficient of relative risk aversion and intertemporal elasticity of substitution are both unity.

When γ in (50) is large, so that $\theta < +\infty$, it means that the type II agent fears model misspecification. Then notice that the risk-aversion term (51) contributes only a small fraction of the total compensation required to accept the certainty equivalent path. Evidently, for the type II agent, the part of the compensation in Eq. (50) that is accounted for by aversion to model uncertainty is[20]

$$\begin{aligned} c_0^{II}(r) - c_0^{II} &= \frac{\beta\sigma_\varepsilon^2}{2(1-\beta)}\left[\frac{1}{(1-\beta)\theta}\right] \\ &= \frac{\beta\sigma_\varepsilon^2}{2(1-\beta)}(\gamma - 1). \end{aligned} \tag{52}$$

8.1.3. Type III agent

Consider next a type III agent with θ chosen to support constraint η on discounted entropy. For the certainty equivalent path (47), the indifference calculation made with value function J given by (44) is

$$\frac{\beta}{(1-\beta)^2}\left(\mu + \frac{\sigma_\varepsilon^2}{2}\right) + \frac{1}{1-\beta}c_0^{III} = \frac{\beta}{(1-\beta)^2}\left[\mu - \frac{\sigma_\varepsilon^2}{(1-\beta)\theta}\right] + \frac{1}{1-\beta}c_0.$$

[19] Of course, this will be the same compensation that a type I agent with $\gamma = 1$ would require.

[20] Here we use the decomposition in logarithms $c_0 - c_0^{II} = (c_0 - c_0^{II}(r)) + (c_0^{II}(r) - c_0^{II})$. Notice that the implied decomposition in the level of consumption $\exp(c_t)$ is multiplicative.

Therefore

$$c_0 - c_0^{III} = \frac{\beta \sigma_\varepsilon^2}{2(1-\beta)}(2\gamma - 1) \tag{53}$$

and

$$c_0^{III}(r) - c_0^{III} = \frac{\beta \sigma_\varepsilon^2}{(1-\beta)}(\gamma - 1). \tag{54}$$

This is twice the compensation (52) required by a type II agent with the same value of θ.

8.1.4. Type IV agent

Finally, consider a type IV agent. Though he ranks plans according to the value function J given by (44), his attitude is really that of a type I agent with $\gamma = 1$ ($\theta = +\infty$) but a pessimistic view of the mean of log consumption growth, with the mean being altered according to

$$\tilde{\mu} = \mu - \frac{\sigma_\varepsilon^2}{(1-\beta)\theta}$$

where the θ in this formula is the robustness parameter of an associated type III agent whose worst-case model our type IV agent believes without doubt. We again obtain (53) when we ask our type IV agent to tell us how much we could lower the certainty equivalent path (47) to render him indifferent between it and the risky path governed by

$$c_{t+1} - c_t = \tilde{\mu} + \sigma_\varepsilon \varepsilon_{t+1}.$$

8.2. Comparison with risky but free-of-model-uncertainty equivalent path

We now describe an alternative measure of the welfare benefits of removing fear of model misspecification. We no longer use the no-risk certainty equivalent path of Section 8.1. We take a different approach. We isolate the compensation for model uncertainty by allowing only one change in the path for consumption, in particular, time 0 consumption. We compare two paths whose risky consumptions for all dates $t \geqslant 1$ are identical, so all compensation for model uncertainty occurs by adjusting time 0 consumption. We adjust c_0 to equate the value functions for (i) a $\theta < +\infty$ type II agent who fears model misspecification with (ii) a $\theta = +\infty$ type II agent who does not fear model misspecification.

Thus, we consider two trajectories for consumption governed by the random walk for log consumption. For both trajectories, we use a common initial condition c_0 to construct identical continuation log consumptions c_t for $t \geqslant 1$. But for the path that liberates the type II agent from fear of model misspecification, we reduce date zero consumption to $c_0^{II}(u)$. For indifference between situations with fear of model misspecification (the left side of the following equation) and without fear of model misspecification (the right side), we require that

$$\frac{\beta}{(1-\beta)^2}\left[\mu - \frac{\sigma_\varepsilon^2}{2(1-\beta)\theta}\right] + \frac{1}{1-\beta}c_0 = \frac{\beta}{(1-\beta)^2}(\mu) + \frac{1}{1-\beta}c_0 + \left(c_0^{II}(u) - c_0\right).$$

In constructing the right side, we have set $\theta = \infty$ and replaced c_0 with $c_0^{II}(u)$. Solving the above equation for $c_0 - c_0^{II}(u)$ gives

Table 3
Benefits of eliminating model risk and uncertainty.

Type	Compensation	Random walk	Trend stationary	Compensation for
I	$c_0 - c_0^I$	$\frac{\beta\sigma_\varepsilon^2}{2(1-\beta)}\gamma$	$\frac{\sigma_\varepsilon^2\beta}{2(1-\beta\rho^2)} + \frac{\beta\sigma_\varepsilon^2(1-\beta)(\gamma-1)}{2(1-\beta\rho)^2}$	risk
II	$c_0 - c_0^{II}$	$\frac{\sigma_\varepsilon^2\beta}{2(1-\beta)} + \frac{\beta\sigma_\varepsilon^2}{2\theta(1-\beta)^2}$	$\frac{\sigma_\varepsilon^2\beta}{2(1-\beta\rho^2)} + \frac{\beta\sigma_\varepsilon^2}{2\theta(1-\beta\rho)^2}$	risk and uncertainty
II	$c_0 - c_0^{II}(r)$	$\frac{\beta\sigma_\varepsilon^2}{2(1-\beta)}$	$\frac{\sigma_\varepsilon^2\beta}{2(1-\beta\rho^2)}$	risk
II	$c_0^{II}(r) - c_0^{II}$	$\frac{\beta\sigma_\varepsilon^2}{2\theta(1-\beta)^2}$	$\frac{\beta\sigma_\varepsilon^2}{2\theta(1-\beta\rho)^2}$	uncertainty
III	$c_0 - c_0^{III}$	$\frac{\beta\sigma_\varepsilon^2}{\theta(1-\beta)^2} + \frac{\beta\sigma_\varepsilon^2}{2(1-\beta)}$	$\frac{\beta\sigma_\varepsilon^2}{\theta(1-\rho\beta)^2} + \frac{\beta\sigma_\varepsilon^2}{2(1-\beta\rho^2)}$	risk and uncertainty
III	$c_0 - c_0^{III}(r)$	$\frac{\beta\sigma_\varepsilon^2}{2(1-\beta)}$	$\frac{\sigma_\varepsilon^2\beta}{2(1-\beta\rho^2)}$	risk
III	$c_0^{III}(r) - c_0^{III}$	$\frac{\beta\sigma_\varepsilon^2}{\theta(1-\beta)^2}$	$\frac{\beta\sigma_\varepsilon^2}{\theta(1-\rho\beta)^2}$	uncertainty

$$\begin{aligned} c_0 - c_0^{II}(u) &= \frac{\beta\sigma_\varepsilon^2}{2(1-\beta)^2}\left[\frac{1}{(1-\beta)\theta}\right] \\ &= \frac{\beta\sigma_\varepsilon^2}{2(1-\beta)^2}(\gamma-1). \end{aligned} \tag{55}$$

Note how this formula is $\frac{1}{1-\beta}$ times the expression on the right side of (52).

Consider next a type III agent facing the same choice, that is, the same compensation scheme for being able to avoid model uncertainty. Let $c_0^{III}(u)$ denote the consumption that leaves the agent indifferent between the risky but uncertain process and the risky *and* uncertain processes. Then

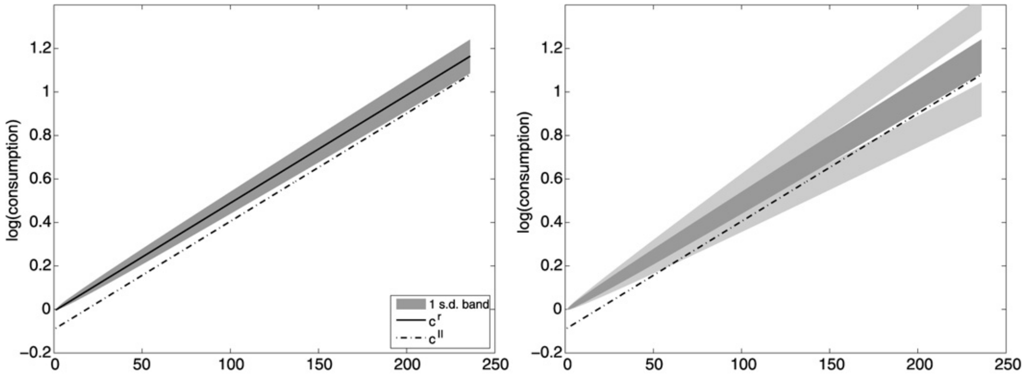
$$\frac{\beta}{(1-\beta)^2}\left[\mu - \frac{\sigma_\varepsilon^2}{(1-\beta)\theta}\right] + \frac{1}{1-\beta}c_0 = \frac{\beta}{(1-\beta)^2}(\mu) + \frac{1}{1-\beta}c_0 + \left(c_0^{III}(u) - c_0\right).$$

Solving for $c_0 - c_0^{III}(u)$ gives

$$\begin{aligned} c_0 - c_0^{III}(u) &= \frac{\beta\sigma_\varepsilon^2}{(1-\beta)^2}\left[\frac{1}{(1-\beta)\theta}\right] \\ &= \frac{\beta\sigma_\varepsilon^2}{(1-\beta)^2}(\gamma-1), \end{aligned}$$

which equals $\frac{1}{1-\beta}$ times the term on the right side of (54).

The compensations in Section 8.1 took a deterministic trajectory as a point of comparison, while the ones here use a random path. Here we exploit the fact that current consumption is known and pile all of the compensation into the first period, leaving the remainder of the path unchanged under the approximating model. Because the only adjustment to consumption occurs at time 0, a multiplicative factor $\frac{1}{1-\beta}$ appears relative to comparable formulas in Section 8.1.



(a) Risk and uncertainty (b) Set of models considered

Fig. 4. Panel A: An elimination of risk and uncertainty experiment for the random walk model. Panel B: Set of models considered by the ambiguity averse agent and an elimination of model uncertainty and risk experiment for the random walk model.

8.3. Formulas for trend stationary model

Table 3 summarizes the above formulas and comparable formulas for the trend stationary model worked out in Appendix A. In the next section, we apply these formulas.

8.4. Quantitative results

The two panels of Fig. 4 are designed to bring out the difference between an elimination-of-risk experiment of the type imagined by Lucas [22,23] and Tallarini [30] and our elimination-of-model-uncertainty experiment. Both panels set $\beta = 0.995$ while calibrating θ to set $p(\theta^{-1}) = 0.10$.

The left panel illustrates our elimination of model uncertainty and risk experiment for a type II agent. The 'fan' in the left panel shows a one-standard deviation band that describes j-step ahead conditional distributions for c for our calibrated random walk model for log consumption. The straight dashed line below the fan shows the certainty equivalent path with date zero consumption reduced by $(c_0 - c_0^{II})$. This reduction makes our representative agent of type II indifferent between this deterministic trajectory and the one illustrated by the 'fan' and therefore compensates him for bearing both risk and model ambiguity. The solid line in the left panel illustrates another certainty equivalent path for a type II consumer who does not fear model uncertainty $(\theta = \infty)$ and therefore measures the contribution from "risk" to $(c_0 - c_0^{II})$. Here the consumption trajectory is initialized at a value $(c_0 - c_0^{II}(r))$ lower than the initial value of the original process. This is the value computed in formula (51). At our calibrated values for the parameters, it is very small. As a result the solid line is only slightly below the center of the 'fan'. So along with Lucas [22,23], we also find that the welfare gains from eliminating well understood risks are very small. We reinterpret the large welfare gains found by Tallarini [30] as coming not from reducing risk, but from reducing model uncertainty.

The right panel illustrates the amount of model uncertainty that our representative agent of type II fears by displaying one standard deviation bands that describe j-step ahead conditional distributions for several stochastic processes: (1) the same calibrated random walk with drift log consumption model depicted in the left panel and (2) two elements drawn from a cloud of models

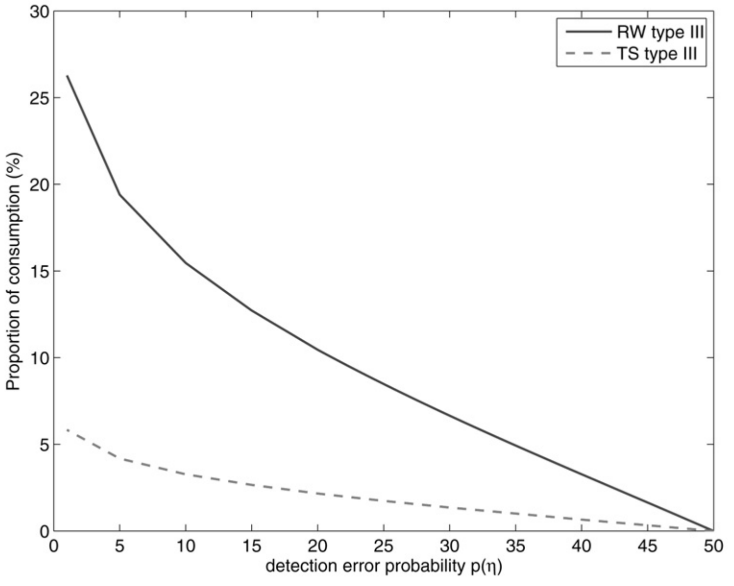


Fig. 5. Proportions $c_0^{III}(r) - c_0^{III}$ of initial consumption that a representative III consumer would surrender not to confront model uncertainty; top line is for random-walk model of consumption growth, bottom line is for trend-stationary model.

that the minimizing player inside our type II representative consumer's head is allowed to choose among, both of which start from the same initial condition as process (1).[21] We also plot the deterministic path (47) that we have initialized using formula (50) to make the agent indifferent between facing this certainty equivalent path and confronting model uncertainty and risk.

As a function of discounted entropy η, Fig. 5 plots the benefits $c_0^{III}(r) - c_0^{III}$ to a type III agent of eliminating model uncertainty as a proportional reduction in initial consumption. The quantities plotted are given by formulas (54) for the random walk model and the corresponding entry in Table 3 for the trend stationary model. The benefits are half as much for a type II agent with a corresponding θ. But relative to the amounts estimated by Lucas [22,23], they are very big.[22]

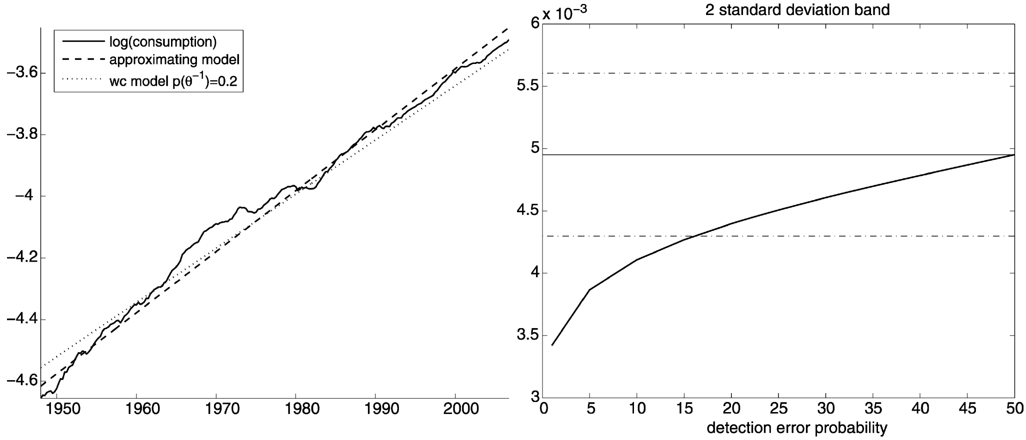
9. Dogmatic Bayesians and learning

9.1. Robust dogmatic Bayesians

Consider the geometric random walk model for consumption. Tallarini [30] follows many other rational expectations researchers in assigning to his representative consumer a dogmatic

[21] The lower element is a worst-case distribution obtained by adjusting the mean to $\mu + \sigma_\varepsilon w(\theta)$ where $w(\theta)$ is the worst case shock for θ that sets $p(\theta^{-1}) = 0.1$, while the upper element adjusts the mean to $\mu - \sigma_\varepsilon w(\theta)$.

[22] Proposition 10 of Cerreia et al. [3] characterizes more uncertainty averse preference relations in terms of pointwise smaller G functions, in their notation. (For our type II agent with multiplier preferences, their G function equals θ times the present value of discounted entropy.) That finding allows us to measure model uncertainty aversion of our agents in terms of alternative θ's for a type II agent and η's for a type III agent. However, it provides no basis for comparing G's across type II and type III agents.



(a) Log consumption (b) Worst case mean consumption growth

Fig. 6. Left panel: log consumption and two lines; right panel: worst case mean consumption growth versus detection error probability.

prior over the growth rate μ and the innovation volatility σ_ε. One way to think about our type II or III representative consumer is that at the end of the day he examines how robust his evaluations are with respect to alternative dogmatic priors over μ. That leads him to price assets like a type IV consumer who has a prior for the mean of log consumption growth concentrated on $\mu + \sigma_\varepsilon w(\theta)$.

The left panel of Fig. 6 compares log consumption to the two lines $c_0 + \mu t$ and $c_0 + (\mu + \sigma_\varepsilon w(\theta))t$ for a θ associated with a detection error probability $p(\theta^{-1}) = 0.2$. These lines are close in the informal eyeball sense that we have shown them to empirically sophisticated friends and upon asking them to tell us which one fits the data best, have received the modal answer 'difficult for me to tell'. The right panel of Fig. 6 conveys the same idea from a different perspective by plotting $\mu + \sigma_\varepsilon w(\theta)$ as a function of the associated detection error probability $p(\theta)$ compared to a two standard deviation band around the maximum likelihood estimate of μ.

The two panels of Fig. 6 and the logic underlying our detection error probabilities indicate that the differences in mean consumption growth that our analysis features are difficult to distinguish with samples of the size, for example, that Tallarini [30] used to estimate mean consumption growth. Our type II representative agent copes with this situation by doing a prior robustness analysis, but we have effectively constrained him to compare priors, all of which are dogmatic. A natural question to ask is: if it is so difficult to learn about μ, would it not make more sense to endow our representative consumer with a non-dogmatic prior over μ?

9.2. Learning?

An affirmative answer to that question is the starting point for Hansen and Sargent [14], who use the analytical framework of Hansen and Sargent [15] to endow a representative consumer with a non-dogmatic prior over μ. They model μ as a hidden Markov state and allow the representative consumer to learn about mean consumption growth as data arrive. But because he does not completely trust the posterior probabilities that emerge from Bayes' law, the representative consumer engages in a worst-case analysis that leads him to slant posterior probabilities pessimistically. By including a hidden state variable that indexes alternative submodels for consumption growth, Hansen and Sargent [14] also study a difficult on-going model selection problem. They posit an associated set of specification doubts that lead the representative consumer to slant posterior probabilities over the submodels pessimistically. These learning problems are sufficiently difficult that the representative consumer is unable to resolve his specification doubts within a sample of the length that Tallarini [30] studied. Robust learning gives rise to countercyclical uncertainty premia because the representative consumer interprets good news about consumption growth as temporary and bad news about consumption growth as permanent.

10. Concluding remarks

It is easy to agree with Lucas that the coefficients γ that Tallarini calibrated to match asset market data are implausibly high when they are interpreted as measures of atemporal risk aversion. Those high γ's become more plausible when we interpret them as measures of the representative consumer's reluctance to face model uncertainty. How we interpret γ has important ramifications about whether risk premia measure (a) the benefits from reducing well understood stochastic aggregate fluctuations, or (b) the benefits of reducing uncertainty about the representative consumer's stochastic specification for consumption growth. We have argued that they measure (b), not (a).

The main point of Lucas [23] was that after one takes into account what has been achieved by using systematic monetary and fiscal policies to smooth aggregate fluctuations in the post WWII U.S., only small *additional* welfare gains can be attained by smoothing transitory shocks further. Under our robustness interpretation, those transitory shocks play a role excluded by Lucas's analysis: by obscuring the consumer's ability to discriminate among alternative models, they put the consumer in a position in which his concerns about model misspecification make him want evaluations of future outcomes that are cautious with respect to a set of plausible statistically nearby models. The process of constructing worst-case scenarios to assist in making those cautious evaluations transforms transient risks into concerns about misspecifications of lower frequency aspects of the representative consumer's approximating model.[23]

Acknowledgments

We thank David Backus, Marc Giannoni, Sydney Ludvigson, Fabio Maccheroni, Massimo Marinacci, Monika Piazzesi, Martin Schneider, Nancy Stokey, Tomasz Strzalecki, and two referees for very helpful comments on earlier drafts.

Appendix A. Formulas for trend stationary model

The worst case mean of ε_{t+1} for the trend stationary model is

$$w(\theta) = \frac{-\sigma_\varepsilon}{(1-\rho\beta)\theta}. \tag{56}$$

Using this, we find that discounted entropy is

[23] The formulas for the worst-case means w_{t+1} in Subsection 6.12 reveal this transformation. In the simple models of this paper, concerns about misspecification translate into (permanent) distortions in the means of shocks. In more general dynamic models, they translate into more richly altered distortions in frequency responses. See Hansen and Sargent [16, Ch. 7].

$$N(x_t) = \frac{\beta\sigma_\varepsilon^2}{2\theta^2(1-\beta)(1-\beta\rho)^2} \tag{57}$$

so the value function for a type III or IV agent is

$$J(x_t) = \frac{\beta\zeta(1-\rho)}{(1-\beta)(1-\beta\rho)} + \frac{\beta\mu}{(1-\beta)^2} - \frac{\sigma_\varepsilon^2\beta}{\theta(1-\beta)(1-\beta\rho)^2} + \frac{\mu\beta(1-\rho)}{(1-\beta\rho)(1-\beta)}t + \frac{1}{1-\beta\rho}c_t \tag{58}$$

and the value function for a type II agent is

$$W(x_t) = \frac{\beta\zeta(1-\rho)}{(1-\beta)(1-\beta\rho)} + \frac{\beta\mu}{(1-\beta)^2} - \frac{\sigma_\varepsilon^2\beta}{2\theta(1-\beta)(1-\beta\rho)^2} + \frac{\mu\beta(1-\rho)}{(1-\beta\rho)(1-\beta)}t + \frac{1}{1-\beta\rho}c_t. \tag{59}$$

The geometric trend stationary model obeys the following difference equation:

$$c_{t+1} = \rho c_t + \zeta(1-\rho) + \rho\mu + \mu(1-\rho)(t+1) + \sigma_\varepsilon\epsilon_{t+1}.$$

We first construct the path of conditional expectations for our original process with fluctuations. Iterating the equation above forward we find that:

$$c_{t+j} = \rho^j c_t + \mu j + (1-\rho)(\zeta+\mu t)\left(1+\rho+\cdots+\rho^{j-1}\right) + \sigma_\varepsilon\left(\epsilon_{t+j} + \rho\epsilon_{t+j-1} + \cdots + \rho^{j-1}\epsilon_{t+1}\right)$$

and therefore

$$\begin{aligned}\log E_t\left[\exp(c_{t+j})\right] &= \rho^j c_t + \mu j + (1-\rho)(\zeta+\mu t)\left(1+\rho+\cdots+\rho^{j-1}\right) + \frac{\sigma_\varepsilon^2}{2}\left(1+\rho^2+\cdots+\rho^{2(j-1)}\right) \\ &= \rho^j c_t + \mu j + (\zeta+\mu t)\left(1-\rho^j\right) + \frac{\sigma_\varepsilon^2}{2}\frac{1-\rho^{2j}}{1-\rho^2}.\end{aligned}$$

We then proceed to compute the value function under the certainty equivalent trajectory

$$\begin{aligned}\tilde{U}(x_t) &= \sum_{j=0}^{\infty}\beta^j\left(\log E_t\left[\exp(c_{t+j})\right]\right) \\ &= \sum_{j=0}^{\infty}\beta^j\left(\rho^j c_t + \mu j + (\zeta+\mu t)\left(1-\rho^j\right) + \frac{\sigma_\varepsilon^2}{2}\frac{1-\rho^{2j}}{1-\rho^2}\right) \\ &= \frac{\zeta\beta(1-\rho)}{(1-\beta)(1-\beta\rho)} + \frac{\mu\beta}{(1-\beta)^2} + \frac{\sigma_\varepsilon^2\beta}{2(1-\beta)(1-\beta\rho^2)} + \frac{\mu\beta(1-\rho)}{(1-\beta)(1-\rho\beta)}t + \frac{c_t}{1-\rho\beta}.\end{aligned}$$

We report in Table 3 the elimination of risk and uncertainty compensations for the trend stationary model that we have computed using the same procedure as for the random walk model. Note that when $\rho = 1$ the compensating variations are identical to the ones for the random walk model.

The alternative Subsection 8.2 welfare measure that compares the risky but free-of-model-uncertainty equivalent path with the original path for the trend stationary model is

$$c_0 - c_0^{II}(u) = \frac{\beta\sigma_\varepsilon^2}{2\theta(1-\beta)(1-\beta\rho)^2} = \frac{\beta\sigma_\varepsilon^2}{2(1-\beta\rho)^2}(\gamma - 1)$$

for a type II agent. The same measure for a type III agent is given by

$$c_0 - c_0^{III}(u) = \frac{\beta\sigma_\varepsilon^2}{\theta(1-\beta)(1-\beta\rho)^2} = \frac{\beta\sigma_\varepsilon^2}{(1-\beta\rho)^2}(\gamma - 1).$$

References

[1] Fernando Alvarez, Urban J. Jermann, Using asset prices to measure the cost of business cycles, J. Polit. Economy 112 (6) (2004) 1223–1256.

[2] E. Anderson, L.P. Hansen, T.J. Sargent, A quartet of semigroups for model specification, robustness, prices of risk, and model detection, J. Europ. Econom. Assoc. 1 (1) (2003) 68–123.

[3] Simone Cerreia, Fabio Maccheroni, Massimo Marinacci, Luigi Montrucchio, Uncertainty averse preferences, Working paper, Collegio Carlo Alberto, 2008.

[4] John H. Cochrane, Where is the market going? Uncertain facts and novel theories, Fed. Reserve Bank Chicago Econ. Perspect. 21 (6) (1997) 3–37.

[5] Jim Dolmas, Risk preferences and the welfare cost of business cycles, Rev. Econ. Dynam. 1 (1998) 646–676.

[6] L. Epstein, S. Zin, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, Econometrica 57 (1989) 937–969.

[7] Larry G. Epstein, Stanley G. Zin, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis, J. Polit. Economy 99 (2) (1991) 263–286.

[8] I. Gilboa, D. Schmeidler, Maxmin expected utility with non-unique prior, J. Math. Econ. 18 (1989) 141–153.

[9] L.P. Hansen, T.J. Sargent, Robust control and model uncertainty, Amer. Econ. Rev. 91 (2001) 60–66.

[10] L.P. Hansen, T.J. Sargent, G.A. Turmuhambetova, N. Williams, Robust control, min-max expected utility, and model misspecification, J. Econ. Theory 128 (2006) 45–90.

[11] Lars Peter Hansen, Ravi Jagannathan, Implications of security market data for models of dynamic economies, J. Polit. Economy 99 (1991) 225–262.

[12] Lars Peter Hansen, Thomas Sargent, Discounted linear exponential quadratic Gaussian control, IEEE Trans. Autom. Control 40 (5) (1995) 968–971.

[13] Lars Peter Hansen, Thomas J. Sargent, Robust estimation and control under commitment, J. Econ. Theory 124 (2) (2005) 258–301.

[14] Lars Peter Hansen, Thomas J. Sargent, Fragile beliefs and the price of model uncertainty, unpublished, 2007.

[15] Lars Peter Hansen, Thomas J. Sargent, Recursive robust estimation and control without commitment, J. Econ. Theory 136 (2007) 1–27.

[16] Lars Peter Hansen, Thomas J. Sargent, Robustness, Princeton University Press, Princeton, New Jersey, 2008.

[17] Lars Peter Hansen, Thomas Sargent, Thomas Tallarini, Robust permanent income and pricing, Rev. Econ. Stud. 66 (1999) 873–907.

[18] Lars Peter Hansen, Thomas J. Sargent, Neng E. Wang, Robust permanent income and pricing with filtering, Macroecon. Dynam. 6 (2002) 40–84.

[19] P. Klibanoff, M. Marinacci, S. Mukerji, A smooth model of decision making under ambiguity, Econometrica 73 (2005) 1849–1892.

[20] Narayana Kocherlakota, Disentangling the coefficient of relative risk aversion from the elasticity of intertemporal substitution: An irrelevance result, J. Finance 45 (1) (1990) 175–190.

[21] D.M. Kreps, E.L. Porteus, Temporal resolution of uncertainty and dynamic choice, Econometrica 46 (1978) 185–200.
[22] Robert E. Lucas Jr., Models of Business Cycles, Basil Blackwell, Oxford, New York, 1987.
[23] Robert E. Lucas Jr., Macroeconomic priorities, American Economic Review, Papers and Proceedings 93 (2003) 1–14.
[24] Fabio Maccheroni, Massimo Marinacci, Aldo Rustichini, Ambiguity aversion, robustness, and the variational representation of preferences, Econometrica 74 (6) (2006) 1447–1498.
[25] Fabio Maccheroni, Massimo Marinacci, Aldo Rustichini, Dynamic variational preferences, J. Econ. Theory 128 (2006) 4–44.
[26] Maurice Obstfeld, Evaluating risky consumption paths: The role of intertemporal substitutability, Europ. Econ. Rev. 38 (1994) 1471–1486.
[27] John W. Pratt, Risk aversion in the small and in the large, Econometrica 32 (1–2) (1964) 122–136.
[28] R.J. Shiller, Consumption, asset prices and economic fluctuations, Carnegie-Rochester Conf. Ser. Public Pol. 17 (1982) 203–238.
[29] Tomasz Strzalecki, Axiomatic Foundations of Multiplier Preferences, Northwestern University, Department of Economics, 2008.
[30] Thomas D. Tallarini, Risk-sensitive real business cycles, J. Monet. Econ. 45 (3) (2000) 507–532.
[31] Phillip Weil, Nonexpected utility in macroeconomics, Quart. J. Econ. 105 (1) (1990) 29–42.