



Recursive robust estimation and control without commitment

Lars Peter Hansen^a, Thomas J. Sargent^{b,*}

^a*University of Chicago, USA*

^b*New York University and Hoover Institution, USA*

Received 2 May 2006; final version received 25 June 2006

Available online 20 June 2007

Abstract

In a Markov decision problem with hidden state variables, a posterior distribution serves as a state variable and Bayes' law under an approximating model gives its law of motion. A decision maker expresses fear that his model is misspecified by surrounding it with a set of alternatives that are nearby when measured by their expected log likelihood ratios (entropies). Martingales represent alternative models. A decision maker constructs a sequence of robust decision rules by pretending that a sequence of minimizing players choose increments to martingales and distortions to the prior over the hidden state. A risk sensitivity operator induces robustness to perturbations of the approximating model conditioned on the hidden state. Another risk sensitivity operator induces robustness to the prior distribution over the hidden state. We use these operators to extend the approach of Hansen and Sargent [Discounted linear exponential quadratic Gaussian control, *IEEE Trans. Automat. Control* 40(5) (1995) 968–971] to problems that contain hidden states.

© 2006 Elsevier Inc. All rights reserved.

JEL classification: C32; C52; D81; D83; D84

Keywords: Robustness; Hidden state Markov chains; Martingales; Risk sensitivity; Decision theory

1. Introduction

This paper studies robust control problems in which the presence of hidden state variables motivates a decision maker to solve a filtering problem. From the standpoint of robustness, hidden states create problems because they give an informational advantage to the hypothetical 'evil agent' who helps the decision maker construct decision rules that are robust to model misspecification.

* Corresponding author.

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

To ensure that decision rules work well even when a decision maker's baseline probability model is flawed, the evil agent must be allowed to alter the dynamics of the hidden states as well as the probabilities assigned to those states. We allow the evil agent to see the hidden states when he explores those alterations.

It is not obvious how the evil agent should exploit his informational advantage in order to help the decision maker design robust decision rules. In [21], we studied a version of this problem when both the decision maker and his evil twin can commit to their decision rules. In contrast, in the present paper, neither agent can commit. Although Hansen and Sargent [21] construct a recursive representation of the equilibrium of the time 0 zero-sum, two-player game that achieves robustness, that equilibrium assumes that the players do not discount increments to entropy. When players discount future utilities, a failure to discount increments to entropy creates time varying decision rules that make concerns about robustness dissipate over time. To allow concerns about robustness to endure and time-invariant decision rules to prevail, this paper studies Markov perfect equilibria of zero-sum games with symmetric discounting of increments both to entropy and to utility.¹ Interestingly, it turns out that this symmetric approach to discounting entropy and utility implies that, in general, worst case beliefs about hidden states are dynamically inconsistent, although beliefs about observables are dynamically consistent.

In more detail, this paper constructs robust decision rules for discounted dynamic programming problems that confront a decision maker with an incentive to evaluate alternative models and to learn about unknown parameters and other hidden state variables.² Extending robust control theory to include these features allows us to approach unsolved problems in positive and normative economics, for example, computing asset evaluations of robust investors and designing robust monetary and fiscal policies. A criticism of our earlier robust control formulations with persistent fears of model misspecification (for example, [20,26,25,2]) is that they precluded learning by not allowing the decision maker to use new information to refine his approximating model and diminish the set of plausible alternatives against which he seeks robustness.³ This paper confronts that criticism by incorporating learning.

The Bayesian literature on model selection and learning has acquired prominent critics. Turning their backs on Bayes' law and the logic of backward induction, distinguished macroeconomists have warned against using experimentation, learning, and model selection to guide decision making.⁴ These skeptics seem to believe that in any concrete application, the details of a model are bound to be subject to misspecifications that will render recommendations to experiment that come from pursuing the logic of Bayes and backward induction unacceptably fragile.⁵ Our methods

¹ Hansen and Sargent [21, p. 286] indicate how part of what we are doing in this paper is to replace $\beta^{-t}\theta$ in the recursive representations there with θ here.

² For examples of such dynamic programming problems, see Jovanic [30], Jovanic and Nyarko [32,33], and Bergemann and Valimaki [5].

³ See Weiland [43].

⁴ Alan Blinder, Robert E. Lucas, and Martin Feldstein have all argued against experimenting to refine model selection. See Cogley et al. [12] for quotations and further discussion.

⁵ See Brock et al. [6,7] for an approach to model selection and decision making that also emphasizes robustness. Cho and Kasa [9,10] study ongoing model selection problems by using mathematical tools that are closely related to the ones used here, but in different and interesting ways. They use entropy based versions of recursive on-going model specification tests to study a model selection problem faced by a large player whose decisions influence aggregate dynamics. They show a tight relationship between the dynamics produced by their model and those that emerge from some models that have been used in the recent literature on adaptive learning, for example, Cho et al. [11] and Kasa [34].

show a decision maker how to experiment even though he distrusts both his models and the prior that he places over them.

We begin by assuming that, through some unspecified process, a decision maker has arrived at an approximating model that fits historical data well. Our notion of a model is broad enough to subsume the possibility that the model is actually a collection of alternative submodels, each of which can contain unknown parameters. More generally, we formulate the decision environment using a hidden state Markov model, and we index a family of probability models by conditioning on the hidden Markov state. Some of the components of the Markov states can be time invariant indicators of submodels and others can be parameters or hidden states that are associated with the submodels and that can vary over time. A conditional model is a probability specification conditioned on a hidden Markov state. Because the decision-maker fears that each member of his family of approximating conditional models is misspecified, he surrounds each of them with a set of unspecified alternative models whose expected log likelihood ratios (i.e., *relative entropies*) are restricted or penalized. The decision maker believes that the data will be generated by an unknown member of one of these sets. When relative entropies are constrained to be small, the decision maker believes that his model is a good approximation. The decision maker wants robustness against these alternatives because, as Anderson et al. [2] emphasize, perturbations with small relative entropies are statistically difficult to distinguish from the approximating model.

Since the decision maker does not know the hidden Markov state, he is also compelled to weight the alternative conditional models. This paper assumes that at each date the appropriate summary of past signals about the hidden state is the decision maker's posterior under the approximating model, just as it is when the decision maker trusts his model. He makes a robust adjustment to these probabilities to accommodate the fact that they were constructed from the approximating model. As we shall see, the decision maker is not required to respect distortions to the distribution of today's hidden state that were implied by his decision making process at earlier dates. Hansen and Sargent [21] studied a closely related decision problem that requires today's decision maker to commit to those prior distortions of the distribution of today's hidden states.

Section 2 formulates a Markov control problem in which a decision maker with a trusted model receives signals about hidden state variables. Subsequent sections view the model of Section 2 as an approximation, use relative entropy to define a cloud of models that are difficult to distinguish from it statistically, and construct a sequence of decision rules that can work well for all of those models. Section 3 uses results of Hansen and Sargent [21] to represent distortions of an approximating model in terms of martingales that are defined on the same probability space as the approximating model. Section 4 then defines two risk-sensitivity operators, T^1 and T^2 , respectively, that are indexed by penalty parameters (θ_1, θ_2) . These two operators are counterparts of related operators in [21], but there they must appear with a common θ . By allowing the two θ 's to differ here, we let the decision maker focus his concerns about robustness on different aspects of his specification. Thus, in Section 5, we use T^1 to adjust continuation values for concerns about model misspecification, conditioned on knowledge of the hidden state. We use T^2 to adjust continuation values for concerns about misspecification of the distribution of the hidden state. We interpret θ_1 and θ_2 as penalties on pertinent entropy terms. Hansen and Sargent [23] specializes our Section 5 recursions to compute robust decision rules for the linear quadratic Gaussian case.

Section 6 relates the special $\theta_1 = \theta_2$ case to a decision problem under commitment that we analyzed in [21]. When $\theta_1 = \theta_2$ and the discount factor is unity, the problem in this paper coincides with an undiscounted version of the problem in [21]. We discuss the dynamic consistency

of worst case beliefs about the hidden state in Sections 6.4 and 6.6. To prepare the machinery needed to construct stochastic discount factors for use in finance and macroeconomics like those mentioned in Section 10, Section 7 describes the worst case distribution over signals. Section 8 interprets our formulation and suggests modifications of it in terms of the multiple priors models of Epstein and Schneider [16,17]. Section 9 relates our formulation to papers about reducing compound lotteries. To illustrate the single-agent decision theory that we have developed in this paper, Section 10 briefly describes a pure-endowment asset pricing model of Hansen and Sargent [22] that prices assets by reading marginal valuations off of the value function of a single representative agent. Hansen and Sargent [22] combine the logarithmic preference specification of Tallarini [42], an endowment specification closely related to one of Bansal and Yaron [3], and special cases of the recursions in [23], which in turn specialize the recursions in this paper. Hansen and Sargent [22] construct the value function for a representative consumer who lives in a pure endowment economy and is unsure about the specifications of two submodels that might govern consumption growth, as well as being unsure about the probability for mixing those submodels. They deduce a multiplicative adjustment to the market price of risk that is contributed by the representative consumer's concerns about robustness. This model illustrates how the representative consumer's concern about misspecification of the probabilities for mixing submodels and also of the conditional means and other hidden components of consumption growth within each submodel can increase the volatility of the stochastic discount factor and enhance what are typically interpreted as risk premia.⁶ Section 11 concludes. Hansen and Sargent [21] contains an extensive account of related literatures. An application to a decision problem with experimentation and learning about multiple submodels appears in [12].

2. A control problem without model uncertainty

For $t \geq 0$, we partition a state vector as $x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}$, where y_t is observed and z_t is not. A vector s_t of observable signals is correlated with the hidden state z_t and is used by the decision maker to form beliefs about the hidden state. Let Z denote a space of admissible unobserved states, \mathcal{Z} a corresponding sigma algebra of subsets of states, and λ a measure on the measurable space of hidden states (Z, \mathcal{Z}) . Let S denote the space of signals, \mathcal{S} a corresponding sigma algebra, and η a measure on the measurable space (S, \mathcal{S}) of signals.

Let $\{\mathcal{S}_t : t \geq 0\}$ denote a filtration, where \mathcal{S}_t is generated by y_0, s_1, \dots, s_t , where we shall assume that s_1, \dots, s_t is generated by system (1)–(3). We can apply Bayes' rule to τ to deduce a density q_t , relative to the measure λ , for z_t conditioned on information \mathcal{S}_t . Let $\{\mathcal{X}_t : t \geq 0\}$ be a larger filtration where \mathcal{X}_t is generated by $x_0, w_1, w_2, \dots, w_t$. The smallest sigma algebra generated by all states for $t \geq 0$ is $\mathcal{X}_\infty \doteq \bigvee_{t \geq 0} \mathcal{X}_t$; the smallest sigma algebra generated by all signals for $t \geq 0$ is $\mathcal{S}_\infty \doteq \bigvee_{t \geq 0} \mathcal{S}_t$. Let A denote a feasible set of actions, which we take to be a Borel set of some finite dimensional Euclidean space, and let \mathcal{A}_t be the set of A -valued random vectors that are \mathcal{S}_t measurable.⁷

⁶ With enough data, the representative consumer might 'learn his way out of' some of the specification concerns including the submodel mixture probabilities and the parameters within submodel. This happens if the parameter estimates converge and the mixture probabilities concentrate on one submodel. But in the model of Hansen and Sargent [22], the mixture probabilities continue fluctuating broadly within the sample for a quarterly record of post WWII U.S. consumption growth rates.

⁷ We could easily allow A to depend on the observable component of the state.

Signals and states are determined by the transition functions

$$y_{t+1} = \pi_y(s_{t+1}, y_t, a_t), \tag{1}$$

$$z_{t+1} = \pi_z(x_t, a_t, w_{t+1}), \tag{2}$$

$$s_{t+1} = \pi_s(x_t, a_t, w_{t+1}), \tag{3}$$

where $\{w_{t+1} : t \geq 0\}$ is an i.i.d. sequence of random variables. Knowledge of y_0 and π_y allows us to construct y_t recursively from signals and actions. The construction of x_t in Eqs. (1)–(2) and the informational constraint on action processes imply that x_t is \mathcal{X}_t measurable and y_t is \mathcal{S}_t measurable. Substituting (3) into (1) gives

$$y_{t+1} = \pi_y[\pi_s(x_t, a_t, w_{t+1}), y_t, a_t] \doteq \bar{\pi}_y(x_t, a_t, w_{t+1}).$$

Eqs. (2) and (3) determine a conditional density $\tau(z_{t+1}, s_{t+1} | x_t, a_t)$ relative to the product measure $\lambda \times \eta$.

As a benchmark, consider the following decision problem under complete confidence in model (1)–(3) but incomplete information about the state:

Problem 2.1.

$$\max_{a_t \in \mathcal{A}_t: t \geq 0} E \left[\sum_{t=0}^T \beta^t U(x_t, a_t) | \mathcal{S}_0 \right], \quad \beta \in (0, 1)$$

subject to (1)–(3).

To make Problem 2.1 recursive, let $*$ denote a next period value and use τ to construct two densities for the signal:

$$\begin{aligned} \kappa(s^* | y_t, z_t, a_t) &\doteq \int \tau(z^*, s^* | y_t, z_t, a_t) d\lambda(z^*) \\ \zeta(s^* | y_t, q_t, a_t) &\doteq \int \kappa(s^* | y_t, z, a_t) q_t(z) d\lambda(z). \end{aligned} \tag{4}$$

By Bayes' rule,

$$q_{t+1}(z^*) = \frac{\int \tau(z^*, s_{t+1} | y_t, z, a_t) q_t(z) d\lambda(z)}{\zeta(s_{t+1} | y_t, q_t, a_t)} \equiv \pi_q(s_{t+1}, y_t, q_t, a_t). \tag{5}$$

In particular applications, π_q can be computed with methods that specialize Bayes' rule (e.g., the Kalman filter or a discrete time version of the Wonham [45] filter).

Take (y_t, q_t) as the state for a recursive formulation of Problem 2.1. The transition laws are (1) and (5). Let $\pi = \begin{bmatrix} \pi_y \\ \pi_q \end{bmatrix}$. Then we can rewrite Problem 2.1 in the alternative form:

Problem 2.2. Choose a sequence of decision rules for a_t as functions of (y_t, q_t) for each $t \geq 0$ that maximize

$$E \left[\sum_{t=0}^T \beta^t U(x_t, a_t) | \mathcal{S}_0 \right]$$

subject to (1), (5), a given density $q_0(z)$, and the density $\kappa(s_{t+1}|y_t, z_t, a_t)$. The Bellman equation for this problem is

$$W(y, q) = \max_{a \in A} \int \left\{ U(x, a) + \beta \int W^*[\pi(s^*, y, q, a)] \kappa(s^*|y, z, a) d\eta(s^*) \right\} q(z) d\lambda(z). \quad (6)$$

In an infinite horizon version of problem 2.2, $W^* = W$.

2.1. Examples

Examples of Problem 2.2 in economics include [30–33,5]. Examples from outside economics appear in [15]. Problems that we are especially interested in are illustrated in the following four examples. (More examples appear in Section 10.)

Example 2.3. Model Uncertainty I: two submodels. Let the hidden state $z \in \{0, 1\}$ index one of two submodels. Let

$$\begin{aligned} y_{t+1} &= s_{t+1}, \\ z_{t+1} &= z_t, \\ s_{t+1} &= \pi_s(y_t, z, a_t, w_{t+1}). \end{aligned} \quad (7)$$

The hidden state is time invariant. The decision maker has prior probability $\text{Prob}(z = 0) = q$. The third equation in (7) depicts two laws of motion. Cogley et al. [12,13] study the value of monetary policy experimentation in a model in which a is an inflation target and $\pi_s(y, z, a, w) = \bar{\pi}_y(y, z, a, w)$ for $z \in \{0, 1\}$ represent two submodels of inflation-unemployment dynamics.

Example 2.4. Model Uncertainty II: a continuum of submodels. The observable state y takes two possible values $\{y_L, y_H\}$. Transition dynamics are still described by (7), but now there is a continuum of models indexed by the hidden state $z \in [0, 1] \times [0, 1]$ that stands for unknown values of two transition probabilities for an observed state variable y . Given z , we can use the third equation of (7) to represent a two state Markov chain that governs the observable state y (see [15]), $P = \begin{bmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{bmatrix}$, where $(p_{11}, p_{22}) = z$. The decision maker has prior distribution $g_{0,1}(p_{11})g_{0,2}(p_{22})$ on z ; $g_{0,1}$ and $g_{0,2}$ are beta distributions.

Example 2.5. Model Uncertainty III: a components model of income dynamics with an unknown fixed effect in labor income. The utility function $U(a_t)$ is a concave function of consumption a_t ; y_{2t} is the level of financial assets, and $y_{1t} = s_t$ is observed labor income. The evolution equations are:

$$\begin{aligned} y_{1,t+1} &= s_{t+1}, \\ y_{2,t+1} &= R[y_{2,t} + y_{1,t} - a_t], \\ z_{1,t+1} &= z_{1,t}, \\ z_{2,t+1} &= \rho z_{2,t} + \sigma_1 w_{1,t+1}, \\ s_{t+1} &= z_{1,t} + z_{2,t} + \sigma_2 w_{2,t+1}, \end{aligned}$$

where $w_{t+1} \sim \mathcal{N}(0, I)$ is an i.i.d. bivariate Gaussian process, $R \leq \beta^{-1}$ is a gross return on financial assets $y_{2,t}$, $|\rho| < 1$, $z_{1,t}$ is one unobserved constant component of labor income, and $z_{2,t}$ is another unobserved serially correlated component of labor income. A decision maker has a prior q_0 over $(z_{1,0}, z_{2,0})$.

Example 2.6. Estimation of drifting coefficients regression model. The utility function $U(x_t, a_t) = -L(z_t - a_t)$, where L is a loss function and a_t is a time- t estimate of the coefficient vector z_t . The evolution equation is

$$\begin{aligned} y_{t+1} &= s_{t+1}, \\ z_{t+1} &= \rho z_t + \sigma_1 w_{1,t+1}, \\ s_{t+1} &= y_t \cdot z_t + \sigma_2 w_{2,t+1}, \end{aligned}$$

where $w_{t+1} \sim \mathcal{N}(0, I)$ and there is a prior $q_0(z)$ on an initial set of coefficients.

2.2. Modified problems that distrust $\kappa(s^*|y, z, a)$ and $q(z)$

This paper studies modifications of Problem 2.2 in which the decision maker wants a decision rule that is robust to possible misspecifications of equations (1)–(2). We express the Bellman equation as (6) and use the decomposition of ζ in (4). Representation (6) focuses the decision maker's concerns on two aspects of the stochastic structure: the conditional distribution of next period's signals $\kappa(s^*|y, z, a)$ and the distribution over this period's value of the hidden state $q(z)$. We propose a recursive formulation of a robust control and estimation problem that allows a decision maker to doubt either or both of these aspects of his stochastic specification. To prepare the way, Section 3 describes how misspecifications of the decision maker's approximating model can be represented in terms of sequences of nonnegative random variables that form martingales under that approximating model.

3. Using martingales to represent model misspecifications

Hansen and Sargent [21] use a nonnegative \mathcal{X}_t -measurable function M_t with $EM_t = 1$ to create a distorted probability measure that is absolutely continuous with respect to the probability measure over \mathcal{X}_t generated by the model (1)–(2). The random variable M_t is a martingale under this probability measure. Using M_t as a Radon–Nikodym derivative generates a distorted measure under which the expectation of a bounded \mathcal{X}_t -measurable random variable W_t is $\tilde{E}W_t \doteq EM_t W_t$. The *entropy* of the distortion at time t conditioned on date zero information is $E(M_t \log M_t | \mathcal{X}_0)$ or $E(M_t \log M_t | \mathcal{S}_0)$.

3.1. Recursive representations of distortions

We often factor a density F_{t+1} for an \mathcal{X}_{t+1} -measurable random variable as $F_{t+1} = F_t f_{t+1}$ where f_{t+1} is a one-step ahead density conditioned on \mathcal{X}_t . It is also useful to factor M_t . Thus, take a nonnegative martingale $\{M_t : t \geq 0\}$ and form

$$m_{t+1} = \begin{cases} \frac{M_{t+1}}{M_t} & \text{if } M_t > 0, \\ 1 & \text{if } M_t = 0. \end{cases}$$

Then $M_{t+1} = m_{t+1}M_t$ and

$$M_t = M_0 \prod_{j=1}^t m_j. \tag{8}$$

The random variable M_0 has unconditional expectation equal to unity. By construction, m_{t+1} has date t conditional expectation equal to unity. For a bounded random variable W_{t+1} that is \mathcal{X}_{t+1} -measurable, the distorted conditional expectation implied by the martingale $\{M_t : t \geq 0\}$ is

$$\frac{E(M_{t+1}W_{t+1}|\mathcal{X}_t)}{E(M_{t+1}|\mathcal{X}_t)} = \frac{E(M_{t+1}W_{t+1}|\mathcal{X}_t)}{M_t} = E(m_{t+1}W_{t+1}|\mathcal{X}_t)$$

provided that $M_t > 0$. We use m_{t+1} to represent distortions of the conditional probability distribution for \mathcal{X}_{t+1} given \mathcal{X}_t . For each $t \geq 0$, construct the space \mathcal{M}_{t+1} of all nonnegative, \mathcal{X}_{t+1} -measurable random variables m_{t+1} for which $E(m_{t+1}|\mathcal{X}_t) = 1$.

The conditional (on \mathcal{X}_t) relative entropy of a nonnegative random variable m_{t+1} in \mathcal{M}_{t+1} is $\varepsilon_t^1(m_{t+1}) \doteq E(m_{t+1} \log m_{t+1}|\mathcal{X}_t)$.

3.2. Distorting likelihoods with hidden information

The random variable M_t is adapted to \mathcal{X}_t and is a likelihood ratio for two probability distributions over \mathcal{X}_t . The \mathcal{S}_t -measurable random variable $G_t = E(M_t|\mathcal{S}_t)$ implies a likelihood ratio for the reduced information set \mathcal{S}_t ; G_t assigns distorted expectations to \mathcal{S}_t -measurable random variables that agree with M_t , and $\{G_t : t \geq 0\}$ is a martingale adapted to $\{\mathcal{S}_t : t \geq 0\}$.

Define the \mathcal{X}_t -measurable random variable h_t by

$$h_t \doteq \begin{cases} \frac{M_t}{E(M_t|\mathcal{S}_t)} & \text{if } E(M_t|\mathcal{S}_t) > 0, \\ 1 & \text{if } E(M_t|\mathcal{S}_t) = 0 \end{cases}$$

and decompose M_t as

$$M_t = h_t G_t. \tag{9}$$

Decompose entropy as

$$\begin{aligned} E(M_t \log M_t|\mathcal{S}_0) &= E[G_t h_t (\log h_t + \log G_t) |\mathcal{S}_0] \\ &= E(G_t h_t \log h_t|\mathcal{S}_0) + E(G_t \log G_t|\mathcal{S}_0) \end{aligned}$$

where we have dropped an h_t from the last term because $E(h_t|\mathcal{S}_t) = 1$ and G_t is \mathcal{S}_t measurable. Define $\varepsilon_t^2(h_t) \doteq E(h_t \log h_t|\mathcal{S}_t)$ as the conditional (on \mathcal{S}_t) relative entropy.

We now have the tools to represent and measure misspecifications of the two components $\kappa(s^*|y, z, a)$ and $q(z)$ in (6). In (9), M_t distorts the probability distribution of \mathcal{X}_t , h_t distorts the probability of \mathcal{X}_t conditioned on \mathcal{S}_t , G_t distorts the probability of \mathcal{S}_t , and m_{t+1} distorts the probability of \mathcal{X}_{t+1} given \mathcal{X}_t . We use multiplication by m_{t+1} to distort κ and multiplication by h_t to distort q . We use $\varepsilon_t^1(m_{t+1})$ to measure m_{t+1} and $\varepsilon_t^2(h_t)$ to measure h_t .

4. Two pairs of operators

This section introduces two pairs of risk-sensitivity operators, (R_t^1, T^1) and (R_t^2, T^2) . In Section 5, we use the T^1 and T^2 operators to define recursions that induce robust decision rules.

4.1. R_t^1 and T^1

For $\theta > 0$, let V_{t+1} be an \mathcal{X}_{t+1} -measurable random variable for which $E \left[\exp \left(-\frac{V_{t+1}}{\theta} \right) \middle| \mathcal{X}_t \right] < \infty$. Then define

$$\begin{aligned} R_t^1(V_{t+1}|\theta) &= \min_{m_{t+1} \in \mathcal{M}_{t+1}} E(m_{t+1} V_{t+1} | \mathcal{X}_t) + \theta \varepsilon_t^1(m_{t+1}) \\ &= -\theta \log E \left[\exp \left(-\frac{V_{t+1}}{\theta} \right) \middle| \mathcal{X}_t \right]. \end{aligned} \tag{10}$$

The minimizing choice of m_{t+1} is

$$m_{t+1}^* = \frac{\exp \left(-\frac{V_{t+1}}{\theta} \right)}{E \left[\exp \left(-\frac{V_{t+1}}{\theta} \right) \middle| \mathcal{X}_t \right]} \tag{11}$$

where the term in the denominator assures that $E m_{t+1}^* | \mathcal{X}_t = 1$.

In the limiting $\theta = \infty$ case, $R_t^1(V_{t+1}|\infty) = E(V_{t+1} | \mathcal{X}_t)$. Notice that this expectation can depend on the hidden state. When $\theta < \infty$, R_t^1 adjusts $E(V_{t+1} | \mathcal{X}_t)$ by using a worst-case belief about the probability distribution of \mathcal{X}_{t+1} conditioned on \mathcal{X}_t that is implied by the twisting factor (11), as well as adding an entropy penalty. When the conditional moment restriction $E \left[\exp \left(-\frac{V_{t+1}}{\theta} \right) \middle| \mathcal{X}_t \right] < \infty$ is not satisfied, we define R_t^1 to be $-\infty$ on the relevant conditioning events.

When the \mathcal{X}_{t+1} -measurable random variable V_{t+1} takes the special form $W(y_{t+1}, q_{t+1}, z_{t+1})$, the $R_t^1(\cdot|\theta)$ operator implies another operator:

$$\begin{aligned} T^1(W|\theta)(y, q, z, a) &= -\theta \log \int \exp \left(-\frac{W[\pi(s^*, y, q, a), z^*]}{\theta} \right) \\ &\quad \times \tau(z^*, s^* | y, z, a) d\lambda(z^*) d\eta(s^*). \end{aligned}$$

The transformation T^1 maps a value function that depends on next period's state (y^*, q^*, z^*) into a risk-adjusted value function that depends on (y, q, z, a) . Associated with this *risk sensitivity adjustment* T^1 is a worst-case distortion in the transition dynamics for the state and signal process. Let ϕ denote a nonnegative density function defined over (z^*, s^*) satisfying

$$\int \phi(z^*, s^*) \tau(z^*, s^* | y, z, a) d\lambda(z^*) d\eta(s^*) = 1. \tag{12}$$

The corresponding entropy measure is:

$$\int \log[\phi(z^*, s^*)] \phi(z^*, s^*) \tau(z^*, s^* | y, z, a) d\lambda(z^*) d\eta(s^*).$$

In our recursive formulation, we think of ϕ as a possibly infinite dimensional control vector (a density function) and consider the minimization problem:

$$\min_{\phi \geq 0} \int (W[\pi(s^*, y, q, a), z^*] + \theta \log[\phi(z^*, s^*)]) \phi(z^*, s^*) \tau(z^*, s^* | y, z, a) d\lambda(z^*) d\eta(s^*)$$

subject to (12). The associated worst-case density conditioned on \mathcal{X}_t is $\phi_t(z^*, s^*)\tau(z^*, s^*|x_t, a_t)$ where

$$\phi_t(z^*, s^*) = \frac{\exp\left(-\frac{V[\pi(s^*, y_t, q_t, a_t), z^*]}{\theta}\right)}{E\left[\exp\left(-\frac{V[\pi(s_{t+1}, y_t, q_t, a_t), z_{t+1}]}{\theta}\right) \middle| \mathcal{X}_t\right]}. \tag{13}$$

4.2. R_t^2 and T^2

For $\theta > 0$, let \hat{W}_t be an \mathcal{X}_t -measurable function for which $E\left[\exp\left(-\frac{\hat{W}_t}{\theta}\right) \middle| \mathcal{S}_t\right] < \infty$. Then define

$$\begin{aligned} R_t^2(\hat{V}_t|\theta) &= \min_{h_t \in \mathcal{H}_t} E(h_t \hat{V}_t | \mathcal{S}_t) + \theta \varepsilon_t^2(h_t) \\ &= -\theta \log E\left[\exp\left(-\frac{\hat{V}_t}{\theta}\right) \middle| \mathcal{S}_t\right]. \end{aligned} \tag{14}$$

where \mathcal{H}_t is the set of all nonnegative \mathcal{X}_t -measurable random variables for which $E(h_t | \mathcal{S}_t) = 1$.

The minimizing choice of h_t is

$$h_t^* = \frac{\exp\left(-\frac{\hat{W}_t}{\theta}\right)}{E\left[\exp\left(-\frac{\hat{W}_t}{\theta}\right) \middle| \mathcal{S}_t\right]},$$

where the term in the denominator assures that $Eh_t^* | \mathcal{S}_t = 1$.

When an \mathcal{X}_t -measurable function has the special form $\hat{W}_t = \hat{V}(y_t, q_t, z_t, a_t)$, (14) implies another operator

$$T^2(\hat{V}|\theta)(y, q, a) = -\theta \log \int \exp\left[-\frac{\hat{V}(y, q, z, a)}{\theta}\right] q(z) d\lambda(z).$$

The associated minimization problem is:

$$\min_{\psi \geq 0} \int \left[\hat{V}(y, q, z, a) + \theta \log \psi(z)\right] \psi(z) q(z) d\lambda(z)$$

subject to (15), where $\psi(z)$ is a relative density that satisfies:

$$\int \psi(z) q(z) d\lambda(z) = 1 \tag{15}$$

and the entropy measure is

$$\int [\log \psi(z)] \psi(z) q(z) d\lambda(z).$$

The optimized density conditioned on \mathcal{S}_t is $\psi_t(z)q_t(z)$, where

$$\psi_t(z) = \frac{\exp\left(-\frac{\hat{V}(y_t, q_t, z, a_t)}{\theta}\right)}{E\left[\exp\left(-\frac{\hat{V}(y_t, q_t, z, a_t)}{\theta}\right) \middle| \mathcal{S}_t\right]}. \tag{16}$$

5. Control problems with model uncertainty

We propose robust control problems that take $q_t(z)$ as the component of the decision maker's state vector that summarizes the history of signals. The decision maker's model includes the law of motion (5) for q (Bayes' law) under the approximating model (1)–(3). Two recursions that generalize Bellman equation (6) express alternative views about the decision maker's fear of misspecification. A first recursion works with value functions that include the hidden state z as a state variable. Let

$$\check{W}(y, q, z) = U(x, a) + E \left\{ \beta \check{W}^*[\pi(s^*, y, q, a), z^*] | x, q \right\}, \quad (17)$$

where the action a solves:

$$W(y, q) = \max_a E \left[U(x, a) + E \left\{ \beta \check{W}^*[\pi(s^*, y, q, a), z^*] | x, q, a \right\} \middle| y, q, a \right]. \quad (18)$$

The value function \check{W} depends on the hidden state z , whereas the value function W in (6) does not. A second recursion modifies the ordinary Bellman equation (6), which we can express as

$$W(y, q) = \max_a E \left[U(x, a) + E \left\{ \beta W^*[\pi(s^*, y, q, a)] | x, q, a \right\} \middle| y, q, a \right]. \quad (19)$$

Although they use different value functions, without concerns about model misspecification, formulations (17)–(19) imply identical control laws. Furthermore, a $W(y, q)$ that satisfies (19) also obeys (18) by virtue of the law of iterated expectations. Because Bellman equation (19) is computationally more convenient, the pair (17)–(18) is not used in the standard problem without a concern for robustness. However, with a concern about robustness, a counterpart to (17)–(18) becomes useful when the decision maker wants to explore distortions of the joint conditional distribution $\tau(s^*, z^* | y, z, a)$.⁸ Distinct formulations emerge from (18) and (19) when we replace the conditional expectation $E(\cdot | y, q, a)$ with $T^2(\cdot | \theta_2)$ and the conditional expectation $E(\cdot | x, q, a)$ with $T^1(\cdot | \theta_1)$. When $\theta_1 = \theta_2 = +\infty$, (17)–(19) lead to value functions and decision rules equivalent to those from (6). When $\theta_1 < +\infty$ and $\theta_2 < +\infty$, recursions (17)–(19) lead to different decision rules because they take different views about the conditional distributions that the malevolent player wants to distort, or equivalently, about the aspects of the stochastic specification in the approximating model against which the decision maker seeks robustness.

5.1. Which conditional distributions to distort?

The approximating model (1)–(3) makes both tomorrow's signal s^* and tomorrow's state z^* functions of x . When tomorrow's value function depends on s^* but not on z^* as in (19), the minimizing player chooses to distort only $\kappa(s^* | y, z, a)$, which amounts to being concerned about misspecifications of the evolution equation (3) for the signal and not (2) for the hidden state. Such a continuation value function imparts no additional incentive to distort the evolution equation (2) of z^* conditioned on s^* and x .⁹ Such a continuation value that depends on s^* but not on z^*

⁸ Another way to express his concerns is that in this case the decision maker fears that (2) and (3) are both misspecified.

⁹ Dependence between (s^*, z^*) conditioned on x under the approximating model means that in the process of distorting s^* conditioned on (x, a) , the minimizing player may indirectly distort the distribution of z^* conditioned on (x, a) . But he does not distort the distribution of z^* conditioned on (s^*, x, a) .

thus imparts concerns about a limited array of distortions that ignore possible misspecification of the z^* evolution (2). Therefore, when we want to direct the maximizing agent's concerns about misspecification onto the conditional distribution $\kappa(s^*|y, z, a)$, we should form a current period value that depends only on the history of the signal and of the observed state. We do this in recursion (23) below.

However, in some situations, we might want to extend the maximizing player's concerns about misspecification to the joint distribution $\tau(z^*, s^*|y, z, a)$ of z^* and s^* . We can do this by making tomorrow's value function for the minimizing player also depend on z^* . In recursions (20)–(21) below, we form a continuation value function that depends on z^* and thereby extend recursions (17), (18) to incorporate concerns about misspecification of (2).

Thus, (20)–(21) below will induce the minimizing player to distort the distribution of z^* conditional on (s^*, x, a) , while the formulation in (23) will not.

5.1.1. Value function depends on (x, q)

By defining a value function that depends on the hidden state, we focus the decision maker's attention on misspecification of the joint conditional distribution $\tau(z^*, s^*|y, z, a)$ of (s^*, z^*) . We modify recursions (17)–(18) by updating a value function according to

$$W(y, q, z) = U(x, a) + T^1 [\beta W^*(y^*, q^*, z^*) | \theta_1] (x, q, a) \quad (20)$$

after choosing an action according to

$$\max_a T^2 \left(U(x, a) + T^1 [\beta W^*(y^*, q^*, z^*) | \theta_1] (x, q, a) \right) | \theta_2 (y, q, a), \quad (21)$$

for $\theta_1 \geq \underline{\theta}_1$, $\theta_2 \geq \underline{\theta}_2(\theta_1)$ for $\underline{\theta}_1, \underline{\theta}_2$ that make the problems well posed.¹⁰ Updating the value function by recursion (20) makes it depend on (x, q) , while using (21) to guide decisions makes actions depend only on the observable state (y, q) . Thus, continuation value \tilde{W} depends on unobserved states, but actions do not. To retain the dependence of the continuation value on z , (20) refrains from using the T^2 transformation when up-dating continuation values.

The fixed point of (20)–(21) is the value function for an infinite horizon problem. For the finite horizon counterpart, we begin with a terminal value function and view the right side of (20) as mapping next period's value function into the current period value function.

5.1.2. Time inconsistency of maximizing player's preferences

In formulation (20)–(21), the current period decision maker acknowledges the dependence on the current hidden state of discounted future returns. For simplicity, suppose that we set $\theta_1 = \infty$. Then $W(y, q, z)$ gives the discounted value of an objective conditioned on the hidden state z . That this hidden state helps predict future signals and future observable state vectors reflects the dependence of this value function on z . This dependence remains when we let $\theta_1 < \infty$, thus activating a concern about model misspecification conditioned on the current period value of the state z . Such dependence is also present in a commitment formulation of the problem discussed in [21]. In the present formulation without commitment, we use recursion (20) to portray a Markov perfect equilibrium of a game in which the date t maximizing decision maker (and his

¹⁰ Limits on θ_1 and θ_2 are typically needed to make the outcomes of the T^1 and T^2 operators be finite.

malevolent companions) take as given the decisions of future maximizing decision makers (and their malevolent companions).¹¹

That the T^2 operator is applied only at the last stage of the backward induction in (20)–(21) renders the preferences of the time 0 agent dynamically inconsistent.¹² The dynamic inconsistency reflects a conflict between the interests of decision makers at different times, one that vanishes when $\beta \rightarrow 1$ and which we now describe.

To explore the preferences implicit in this formulation it is convenient to apply the operators R_t^1 and R_t^2 to continuation values. Let V_{t+1} denote the continuation values of a stochastic process of actions from date $t + 1$ forward. This continuation value can depend on the future states. It is \mathcal{X}_{t+1} measurable but not necessarily \mathcal{S}_{t+1} measurable. Assess this action process at date $t+1$ using $R_{t+1}^2(V_{t+1}|\theta_2)$, which makes a robust adjustment and results in an \mathcal{S}_{t+1} measurable continuation value.

Consider two such continuation values, V_{t+1}^a and V_{t+1}^b , where

$$R_{t+1}^2(V_{t+1}^a|\theta_2) \geq R_{t+1}^2(V_{t+1}^b|\theta_2). \tag{22}$$

We are interested in a date t ranking of these two after we discount (22) and add a common current period contribution U_t to both before applying R_t^2 . This results in two continuation values that are not necessarily comparable, namely, $U_t + R_t^1(\beta V_{t+1}^a|\theta_1)$ and $U_t + R_t^1(\beta V_{t+1}^b|\theta_1)$. For some realized signal histories, the ranking in inequality (22) can be reversed, even after applying R_t^2 .

It is instructive to consider the special case in which U_t is \mathcal{S}_t measurable. Then

$$R_t^2[U_t + R_t^1(\beta V_{t+1}^j|\theta_1)|\theta_2] = U_t + R_t^2[R_{t+1}^1(\beta V_{t+1}^j|\theta_1)|\theta_2]$$

for $j = a, b$. The source of possible intertemporal reversal of rankings is that inequality (22) does not imply:

$$R_t^2[R_{t+1}^1(\beta V_{t+1}^a|\theta_1)|\theta_2] \geq R_t^2[R_{t+1}^1(\beta V_{t+1}^b|\theta_1)|\theta_2].$$

If, however, we strengthen inequality (22) to be

$$V_{t+1}^a \geq V_{t+1}^b$$

then the rankings are preserved. Thus, when we limit comparisons to ones conditioned on hidden states, then intertemporal inconsistency vanishes.

In the next subsection, we propose an alternative approach that avoids the conflict that is the source of this intertemporal inconsistency at the cost of giving the hidden states a less direct role. In particular, as we shall see, this alternative approach considers value functions that depend only on (y, q) and not the hidden state z . This formulation removes an incentive to explore misspecification of the hidden state dynamics themselves and instead focuses only on how those misspecifications might affect the evolution of signals.

5.2. Value function depends on (y, q)

To focus on misspecifications of the conditional distribution $\kappa(s^*|y, z, a)$, we want the minimizing player’s value function to depend only on the reduced information encoded in (y, q) .

¹¹ Laibson [37] uses a Markov perfect equilibrium of such a game to model the decisions of made by someone with intertemporally inconsistent preferences coming from hyperbolic discounting.

¹² That dynamic inconsistency is what prompts us to model decisions as the Markov perfect equilibrium represented in recursion (20).

For this purpose, we use the following counterpart to recursion (19):

$$W(y, q) = \max_a T^2 \left(U(x, a) + T^1 [\beta W^*(y^*, q^*) | \theta_1] (x, q, a) \middle| \theta_2 \right) (y, q, a) \tag{23}$$

for $\theta_1 \geq \underline{\theta}_1$ and $\theta_2 \geq \underline{\theta}_2(\theta_1)$. Although z^* is excluded from the value function W^* , z may help predict the observable state y^* or it may enter directly into the current period reward function, so application of the operator T^1 creates a value function that depends on (x, q, a) , including the hidden state z . Since the malevolent agent observes z , he can distort the dynamics for the observable state conditioned on z via the T^1 operator. Subsequent application of T^2 gives a value function that depends on (y, q, a) , but not z ; T^2 distorts the hidden state distribution. The decision rule sets action a as a function of (y, q) . The fixed point of Bellman equation (23) gives the value function for an infinite horizon problem. For finite horizon problems, we iterate on the mapping defined by the right side of (23), beginning with a known terminal value function. Recursion (23) extends the recursive formulation of risk-sensitivity with discounting advocated by Hansen and Sargent [20] to situations with a hidden state.

5.3. A third formulation that forgets that z is hidden

It is interesting to contrast the above approaches with an alternative one that is feasible for problems in which z does not appear directly in U but instead either y appears alone or y and q both appear. Then one could proceed by simply applying a single risk-sensitivity operator. For such problems, the Bellman equation without concerns about robustness (6) could also be expressed as

$$W(y, q) = \max_{a \in A} U(y, q, a) + \beta \int W^* [\pi(s^*, y, q, a)] \zeta(s^* | y, q, a) d\eta(s^*).$$

The analysis of robust control problems without hidden states in [20,26] could be applied to obtain robust decision rules by taking (y, q) as the observed state. Decision rules that are robust to misspecification of $\zeta(s^* | y, q, a)$ can be obtained by iterating on

$$W(y, q) = \max_{a \in A} U(y, q, a) + T^1 [\beta W^*(y^*, q^*) | \theta](y, q, a).$$

This approach absorbs Bayes' law into the transition law for the state and seeks robustness to misspecification of $\zeta(s^* | y, q, a)$. In contrast, the formulations in (20)–(21) and (23) distinguish distortions to $\kappa(s^* | y, z, a)$ and to $q(z)$ and seek robustness to misspecifications of each of them separately.

5.4. Advantages of our specification

We take the distribution $q_t(z)$ as a state variable and explore misspecifications of it. An alternative way to describe a decision maker's fears of misspecification would be to perturb the evolution equation for the hidden state (2) directly. Doing that would complicate the problem substantially by requiring us to solve a filtering problem for each perturbation of (2). Our formulation avoids multiple filtering problems by solving one and only one filtering problem under the approximating model. The transition law π_q for $q(z)$ in (5) becomes a component of the approximating model.

When $\theta_1 = +\infty$ but $\theta_2 < +\infty$, the decision maker trusts the signal dynamics $\kappa(s^* | y, z, a)$ but distrusts $q(z)$. When $\theta_2 = +\infty$ but $\theta_1 < +\infty$, the situation is reversed. The two- θ formulation

thus allows the decision maker to disentangle his suspicions about these two aspects of the model. Before saying more about the two- θ formulation, the next section explores some ramifications of the special case in which $\theta_1 = \theta_2$ and how it compares to the single θ specification that prevails in a related decision problem under commitment.

6. The $\theta_1 = \theta_2$ case

For the purpose of studying intertemporal consistency and other features of the associated worst case models, it is interesting to compare the outcomes of recursions (20)–(21) or (23) with the decision rule and worst case model described by Hansen and Sargent [21] in which at time 0 the maximizing and minimizing players in a zero-sum game commit to a sequence of decision rules and a single worst case model, respectively. Because there is a single robustness parameter θ in this “commitment model”, it is natural to make this comparison for the special case in which $\theta_1 = \theta_2$.

6.1. A composite operator $T^2 \circ T^1$ when $\theta_1 = \theta_2$

When a common value of θ appears in the two operators, the sequential application $T^2 T^1$ can be replaced by a single operator:

$$\begin{aligned} T^2 \circ T^1 [U(x, a) + \beta W(y^*, q^*)] (y, q, a) \\ = -\theta \log \int \exp \left(-\frac{U(x, a) + \beta W[\pi(s^*, y, q, a)]}{\theta} \right) \kappa(s^* | y, z, a) q(z) d\eta(s^*) d\lambda(z). \end{aligned}$$

This operator is the outcome of a portmanteau minimization problem over a single relative density $\varphi(s^*, z) \geq 0$ that satisfies¹³

$$\int \varphi(s^*, z) \kappa(s^* | y, z, a) q(z) d\eta(s^*) d\lambda(z) = 1,$$

where φ is related to ϕ and ψ defined in (12) and (15) by

$$\varphi(s^*, z) = \int \phi(z^*, s^* | z) \psi(z) q^*(z^*) d\lambda(z^*),$$

where this notation emphasizes that the choice of ϕ can depend on z . The entropy measure for φ is

$$\int [\log \varphi(s^*, z)] \varphi(s^*, z) \kappa(s^* | y, z, a) q(z) d\eta(s^*) d\lambda(z),$$

and the minimizing composite distortion φ to the joint density of (s^*, z) given \mathcal{S}_t is

$$\varphi_t(s^*, z) = \frac{\exp \left(-\frac{U(y_t, z, a_t) + \beta W[\pi(s^*, y_t, q_t, a_t)]}{\theta} \right)}{E \left[\exp \left(-\frac{U(y_t, z, a_t) + \beta W[\pi(s_{t+1}, y_t, q_t, a_t)]}{\theta} \right) \middle| \mathcal{S}_t \right]}. \tag{24}$$

¹³ Recall that applying T^1 and T^2 separately amounts to minimizing over separate relative densities ϕ and ψ .

6.2. Special case $U(x, a) = \hat{U}(y, a)$

When $U(x, a) = \hat{U}(y, a)$, the current period utility drops out of formula (24) for the worst-case distortion to the distribution, and it suffices to integrate with respect to the distribution $\varsigma(s^*|y, q, a)$ that we constructed in (4) by averaging κ over the distribution of the hidden state. Probabilities of future signals compounded by the hidden state are simply averaged out using the state density under the benchmark model, a reduction of a compound lottery that would not be possible if different values of θ were to occur in the two operators.

To understand these claims, we deduce a useful representation of $\varepsilon_t(m_{t+1}, h_t)$ by solving:

$$\varepsilon_t(m_{t+1}^*, h_t^*) \equiv \min_{m_{t+1} \in \mathcal{M}_t, h_t \in \mathcal{H}_t} E \left[h_t \varepsilon_t^1(m_{t+1}) | \mathcal{S}_t \right] + \varepsilon_t^2(h_t)$$

subject to $E(m_{t+1} h_t | \mathcal{S}_{t+1}) = g_{t+1}$, where $E(g_{t+1} | \mathcal{S}_t) = 1$, a constraint that we impose because our aim is to distort expectations of \mathcal{S}_{t+1} -measurable random variables given current information \mathcal{S}_t . The minimizers are

$$m_{t+1}^* = \begin{cases} \frac{g_{t+1}}{E(g_{t+1} | \mathcal{X}_t)} & \text{if } E(g_{t+1} | \mathcal{X}_t) > 0, \\ 1 & \text{if } E(g_{t+1} | \mathcal{X}_t) = 0 \end{cases}$$

and $h_t^* = E(g_{t+1} | \mathcal{X}_t)$. Therefore, $m_{t+1}^* h_t^* = g_{t+1}$ and the minimized value of the objective is

$$\varepsilon_t(m_{t+1}^*, h_t^*) = E \left[g_{t+1} \log(g_{t+1}) | \mathcal{S}_t \right] \equiv \tilde{\varepsilon}_t(g_{t+1}). \tag{25}$$

Thus, in penalizing distortions to continuation values that are \mathcal{S}_t -measurable, it suffices to use the entropy measure $\tilde{\varepsilon}_t$ defined in (25) and to explore distortions to the conditional probability of \mathcal{S}_{t+1} -measurable events given \mathcal{S}_t . This is precisely what the g_{t+1} random variable accomplishes. The g_{t+1} associated with $\mathbb{T}^2 \mathbb{T}^1$ in the special case in which $U(x, a) = \hat{U}(y, a)$ implies a distortion ϕ_t in Eq. (13) that depends on s^* alone. The iterated operator $\mathbb{T}^2 \mathbb{T}^1$ can be regarded as a single risk-sensitivity operator that functions like \mathbb{T}^1 :

$$\begin{aligned} & \mathbb{T}^2 \mathbb{T}^1 \left[\hat{U}(y, a) + \beta W^*(y^*, q^*) \right] (y, q, a) \\ &= \hat{U}(y, a) - \theta \log \int \exp \left(- \frac{\beta W^*(\pi(s^*, y, q, a))}{\theta} \right) \varsigma(s^* | y, q, a) d\eta(s^*). \end{aligned}$$

In [23], we describe how to compute this operator for linear quadratic problems.

6.3. Role of absolute continuity and relation to commitment solutions

Among the outcomes of iterations on the recursions (20)–(21) or (23) of Section 5 are time-invariant functions that map (y_t, q_t) into a pair of nonnegative random variables (m_{t+1}, h_t) . For the moment, ignore the distortion h_t and focus exclusively on m_{t+1} . Through (8), the time-invariant rule for m_{t+1} can be used to construct a martingale $\{M_t : t \geq 0\}$. This martingale implies a limiting probability measure on $\mathcal{X}_\infty = \bigvee_{t \geq 0} \mathcal{X}_t$ via the Kolmogorov extension theorem. The implied probability measure on \mathcal{X}_∞ will typically not be absolutely continuous over the entire collection of limiting events in \mathcal{X}_∞ . Although the martingale converges almost surely by virtue of Doob’s martingale convergence theorem, without absolute continuity, the limiting random variable will not have unit expectation. This implies that concerns about robustness persist in a

way that they don't in a class of robust control problems under commitment that are studied, for example, by Whittle [44] and Hansen and Sargent [21].¹⁴

6.3.1. Problem formulation

Let M_∞ be a nonnegative random variable that is measurable with respect to \mathcal{X}_∞ , with $E(M_\infty | \mathcal{S}_0) = 1$. For a given action process $\{a_t : t \geq 0\}$ adapted to $\{\mathcal{X}_t : t \geq 0\}$, let $V_\infty \doteq \sum_{t=0}^\infty \beta^t U(x_t, a_t)$ subject to (1)–(2). Suppose that $\theta > 0$ is such that $E[\exp(-\frac{1}{\theta} V_\infty) | \mathcal{S}_0] < \infty$. Then

$$R_\infty^1(V_\infty) \doteq \min_{M_\infty \geq 0, E(M_\infty | \mathcal{S}_0) = 1} E(M_\infty V_\infty | \mathcal{S}_0) + \theta E(M_\infty \log M_\infty | \mathcal{S}_0) \tag{27}$$

$$= -\theta \log E\left[\exp\left(-\frac{1}{\theta} V_\infty\right) \middle| \mathcal{S}_0\right]. \tag{28}$$

This static problem has minimizer $M_\infty^* = \frac{\exp(-\frac{1}{\theta} V_\infty)}{E[\exp(-\frac{1}{\theta} V_\infty) | \mathcal{S}_0]}$ that implies a martingale $M_t^* = E(M_\infty^* | \mathcal{X}_t)$.¹⁵ Control theory interprets (28) as a risk-sensitive adjustment of the criterion V_∞ (e.g., see [44]) and gets decisions that are robust to misspecifications by solving

$$\max_{a_t \in \mathcal{A}_t, t \geq 0} -\theta \log E\left[\exp\left(-\frac{1}{\theta} V_\infty\right) \middle| \mathcal{S}_0\right].$$

In a closely related setting, Whittle [44] obtained time-varying decision rules for a_t that converge to ones that ignore concerns about robustness (i.e., those computed with $\theta = +\infty$).

The dissipation of concerns about robustness in this commitment problem is attributable to setting $\beta \in (0, 1)$ while using the undiscounted form of entropy in the criterion function (27). Those features lead to the existence of a well defined limiting random variable M_∞ with expectation unity (conditioned on \mathcal{S}_0), which means that tail events that are assigned probability zero under the approximating model are also assigned probability zero under the distorted model.¹⁶

¹⁴ The product decomposition (8) of M_t implies an additive decomposition of entropy:

$$E(M_t \log M_t | \mathcal{S}_0) - E(M_0 \log M_0 | \mathcal{S}_0) = \sum_{j=0}^{t-1} E[M_j E(m_{j+1} \log m_{j+1} | \mathcal{X}_j) | \mathcal{S}_0]. \tag{26}$$

Setting $E(M_0 | \mathcal{S}_0) = 1$ means that we distort probabilities conditioned on \mathcal{S}_0 .

¹⁵ See Dupuis and Ellis [14]. While robust control problems are often formulated as deterministic problems, here we follow Petersen et al. [40] by studying a stochastic version with a relative entropy penalty.

¹⁶ Because all terms on the right side of (26) are nonnegative, the sequence

$$\sum_{j=0}^{t-1} M_{j-1} E(m_j \log m_j | \mathcal{X}_{j-1})$$

is increasing. Therefore, it has a limit that might be $+\infty$ with positive probability. Thus, $\lim_{t \rightarrow \infty} E(M_t \log M_t | \mathcal{S}_0)$ converges. Hansen and Sargent [21] show that when this limit is finite almost surely, the martingale sequence $\{M_t : t \geq 0\}$ converges in the sense that $\lim_{t \rightarrow \infty} E(|M_t - M_\infty| | \mathcal{S}_0) = 0$, where M_∞ is measurable with respect to $\mathcal{X}_\infty \doteq \bigvee_{t=0}^\infty \mathcal{X}_t$. The limiting random variable M_∞ can be used to construct a probability measure on \mathcal{X}_∞ that is absolutely continuous with respect to the probability measure associated with the approximating model. Moreover, $M_t = E(M_\infty | \mathcal{X}_t)$. When the implied M_∞ is strictly positive with probability one, the distorted probability measure will be equivalent with the original probability measure. In this case, tail events that are assigned probability measure zero under either measure are assigned zero under the other one.

6.3.2. Persistence of robustness concerns without commitment

In our recursive formulations (20)–(21) and (23) of Section 5, the failure of the worst-case nonnegative martingale $\{M_t : t \geq 0\}$ to converge to a limit with expectation one (conditioned on \mathcal{S}_0) implies that the distorted probability distribution on \mathcal{X}_∞ is not absolutely continuous with respect to the probability distribution associated with the approximating model. This feature sustains enduring concerns about robustness and permits time-invariant robust decision rules, in contrast to the outcomes with discounting in [44,21], for example. For settings with a fully observed state vector, Hansen and Sargent [20] and Hansen et al. [26] discounted entropy in order to formulate recursive problems that yield time-invariant decision rules and enduring concerns about robustness. The present paper extends these recursive formulations to problems with unobserved states.

6.4. Dynamic inconsistency of worst-case probabilities about hidden states

This section links robust control theory to recursive models of uncertainty aversion by exploring aspects of the worst case probability models that emerge from the recursions defined in Section 5. Except in a special case that we describe in Section 6.6, those recursions achieve dynamic consistency of decisions by sacrificing dynamic consistency of beliefs about hidden state variables. We explore how this happens. Until we get to the special case analyzed in Section 6.6, the arguments of this subsection will also apply to the general case in which $\theta_1 \neq \theta_2$.

Problems (10) and (14) that define R_t^1 and R_t^2 , respectively, imply worst-case probability distributions that we express as a pair of Radon–Nikodym derivatives (m_{t+1}^*, h_t^*) . Are these probability distortions consistent with next period’s distortion h_{t+1}^* ? Not necessarily, because we have not imposed the pertinent consistency condition on these beliefs. In particular, our use of m_{t+1}, h_t to distort two conditional distributions each period overdetermines a distortion to the distribution of x_{t+1} conditional on \mathcal{S}_{t+1} : because m_{t+1} distorts the probabilities of \mathcal{X}_{t+1} events conditional on \mathcal{X}_t and h_t distorts the probabilities of \mathcal{X}_t events conditioned on \mathcal{S}_t , $m_{t+1}h_t$ distorts the probabilities of \mathcal{X}_{t+1} events conditional on \mathcal{S}_t . Given the distorted probabilities of \mathcal{X}_{t+1} events conditioned on \mathcal{S}_t , we can deduce the probability distortion of \mathcal{X}_{t+1} events conditional on \mathcal{S}_{t+1} (because $\mathcal{S}_t \subset \mathcal{S}_{t+1} \subset \mathcal{X}_{t+1}$). If we had required the decision maker at time $t + 1$ to adhere to this distortion, he would not be free to choose h_{t+1} anew at time $t + 1$. Thus, except when a special condition that we lay out in the next subsection is met, the decision maker’s worst-case beliefs about the distribution of x_{t+1} conditional on \mathcal{S}_{t+1} will not be time-consistent. This is a price that we pay to attain a recursive formulation in which $q_t(z)$ remains a state variable for our formulation of the robust estimation and control problem.

6.5. A belief consistency condition

To deduce a sufficient condition for time consistency, recall that the implied $\{M_{t+1}^* : t \geq 0\}$ should be a martingale. Decompose M_{t+1}^* in two ways:

$$M_{t+1}^* = m_{t+1}^* h_t^* G_t^* = h_{t+1}^* G_{t+1}^*.$$

These equations involve G_{t+1}^* and G_t^* , both of which we have ignored in the recursive formulation of Section 5. Taking expectations conditioned on \mathcal{S}_{t+1} on both sides of $m_{t+1}^* h_t^* G_t^* = h_{t+1}^* G_{t+1}^*$ yields

$$G_t^* E(m_{t+1}^* h_t^* | \mathcal{S}_{t+1}) = G_{t+1}^*.$$

Thus,

$$g_{t+1}^* = E(m_{t+1}^* h_t^* | \mathcal{S}_{t+1})$$

is the implied multiplicative increment for the candidate martingale $\{G_t^* : t \geq 0\}$ adapted to the signal filtration.

Claim 6.1. *A sufficient condition for the distorted beliefs to be time consistent is that the process $\{h_t^* : t \geq 0\}$ should satisfy:*

$$h_{t+1}^* = \begin{cases} \frac{m_{t+1}^* h_t^*}{E(m_{t+1}^* h_t^* | \mathcal{S}_{t+1})} & \text{if } E(m_{t+1}^* h_t^* | \mathcal{S}_{t+1}) > 0, \\ 1 & \text{if } E(m_{t+1}^* h_t^* | \mathcal{S}_{t+1}) = 0. \end{cases} \tag{29}$$

This condition is necessary if $G_{t+1}^* > 0$.¹⁷

The robust control problem under commitment analyzed by Hansen and Sargent [21] satisfies condition (29) by construction: at time 0 a single minimizing player chooses a pair (m_{t+1}^*, h_t^*) that implies next period's h_{t+1}^* . However, in the recursive games defined in the recursions (20)–(21) and (23) in Section 5, the date $t + 1$ minimizing agent can deviate from the h_{t+1}^* that is implied by the (m_{t+1}^*, h_t^*) pair chosen by the date t minimizing agent. The pair (m_{t+1}^*, h_t^*) gives one distortion of the distribution of the hidden state (conditioned on \mathcal{S}_{t+1}) and h_{t+1}^* gives another. We do not require that these agree, and, in particular, do not require that the probabilities of events in \mathcal{X}_{t+1} be distorted in the same ways by the date t determined worst-case distribution (conditioned on \mathcal{S}_{t+1}) and the date $t + 1$ worst-case distribution (conditioned on \mathcal{S}_{t+1}).

A conflict can arise between these worst-case distributions because choosing an action is forward-looking, while estimation of z is backward looking. Dynamic inconsistency of any kind is a symptom of conflicts among the interests of different decision makers, and that is the case here. The two-player games that define the evaluation of future prospects (T^1) and estimation of the current position of the system (T^2) embody different orientations— T^1 looking to the future, T^2 focusing on an historical record of signals.

The inconsistency of the worst-case beliefs pertains only to the decision maker's opinions about the hidden state. If we ignore hidden states and focus on signals, we can assemble a consistent distorted signal distribution by constructing $g_{t+1}^* = E(m_{t+1}^* h_t^* | \mathcal{S}_{t+1})$ and noting that $E(g_{t+1}^* | \mathcal{S}_t) = 1$, so that g_{t+1}^* is the implied one-period distortion in the signal distribution. We can construct a distorted probability distribution over events in \mathcal{S}_{t+1} by using

$$G_{t+1}^* = \prod_{j=1}^{t+1} g_j^*. \tag{30}$$

¹⁷This consistency condition arguably could be relaxed for the two player game underlying (23). Although we allow m_{t+1} to depend on the signal s_{t+1} and the hidden state z_{t+1} , the minimizing solution associated with recursions (23) depends only on the signal s_{t+1} . Thus we could instead constrain the minimizing agent in his or her choice of m_{t+1} and introduce a random variable \tilde{m}_{t+1} that distorts the probability distribution of z_{t+1} conditioned on s_{t+1} and \mathcal{X}_t . A weaker consistency requirement is

$$h_{t+1}^* = \frac{\tilde{m}_{t+1} m_{t+1}^* h_t^*}{E(\tilde{m}_{t+1} m_{t+1}^* h_t^* | \mathcal{S}_{t+1})}$$

for some \tilde{m}_{t+1} with expectation equal to one conditioned on s_{t+1} and \mathcal{X}_t .

Under this interpretation, the pair (m_{t+1}^*, h_t^*) is only a device to construct g_{t+1}^* . When the objective function U does not depend directly on the hidden state vector z , as is true in many economic problems, the consistent set of distorted probabilities defined by (30) describes the events that directly influence the decision maker's well being.

6.6. Discounting and payoffs influenced by hidden states are the source of intertemporal inconsistency

If $\beta = 1$ and $U(x, a)$ does not depend on the hidden state, we can show that the distortions (m_{t+1}, h_t) implied by our recursions satisfy the restriction required for Claim 6.1 and so are temporally consistent. Therefore, in this special case, the recursive games in Section 5 imply the same decisions and worst case distortions as the game under commitment analyzed by Hansen and Sargent [21]. For simplicity, suppose that we fix an action process $\{a_t : t \geq 0\}$ and focus exclusively on assigning distorted probabilities. Let $\{V_t : t \geq 0\}$ denote the process of continuation values determined recursively and supported by choices of worst-case models.

Consider two operators R_t^1 and R_t^2 with a common θ . The operator R_t^1 implies a worst-case distribution for \mathcal{X}_{t+1} conditioned on \mathcal{X}_t with density distortion:

$$m_{t+1}^* = \frac{\exp\left(-\frac{V_{t+1}}{\theta}\right)}{E\left[\exp\left(-\frac{V_{t+1}}{\theta}\right) \middle| \mathcal{X}_t\right]}.$$

The operator R_t^2 implies a worst-case model for the probability of \mathcal{X}_t conditioned on \mathcal{S}_t with density distortion:

$$h_t^* = \frac{E\left[\exp\left(-\frac{V_{t+1}}{\theta}\right) \middle| \mathcal{X}_t\right]}{E\left[\exp\left(-\frac{V_{t+1}}{\theta}\right) \middle| \mathcal{S}_t\right]}.$$

Combining the distortions gives

$$m_{t+1}^* h_t^* = \frac{\exp\left(-\frac{V_{t+1}}{\theta}\right)}{E\left[\exp\left(-\frac{V_{t+1}}{\theta}\right) \middle| \mathcal{S}_t\right]}.$$

To establish temporal consistency, from Claim 6.1 we must show that

$$h_{t+1}^* = \frac{\exp\left(-\frac{V_{t+1}}{\theta}\right)}{E\left[\exp\left(-\frac{V_{t+1}}{\theta}\right) \middle| \mathcal{S}_{t+1}\right]},$$

where

$$h_{t+1}^* \doteq \frac{E\left[\exp\left(-\frac{V_{t+2}}{\theta}\right) \middle| \mathcal{X}_{t+1}\right]}{E\left[\exp\left(-\frac{V_{t+2}}{\theta}\right) \middle| \mathcal{S}_{t+1}\right]}.$$

This relation is true when $\beta = 1$ and U does not depend on the hidden state z . To accommodate $\beta = 1$, we shift from an infinite horizon problem to a finite horizon problem with a terminal value

function. From value recursion (20) and the representation of R_{t+1}^1 in (10),

$$\exp\left(-\frac{V_{t+1}}{\theta}\right) \propto E\left[\exp\left(-\frac{V_{t+2}}{\theta}\right) \middle| \mathcal{X}_{t+1}\right],$$

where the proportionality factor is \mathcal{S}_{t+1} measurable. The consistency requirement for h_{t+1}^* is therefore satisfied.

The preceding argument isolates the role that discounting plays in rendering the worst case beliefs over the hidden state time inconsistent. Heuristically, the games defined by the recursions (20)–(21) or (23) imply intertemporal inconsistency when $\beta < 1$ because the decision maker discounts both current period returns and current period increments to entropy; while in the commitment problem analyzed in [21], the decision maker discounts current period returns but not current period increments to entropy.

7. Implied worst case model of signal distortion

The martingale (relative to \mathcal{S}_t) increment $g_{t+1} = E(m_{t+1}h_t|\mathcal{S}_t)$ distorts the distribution of the date $t + 1$ signal given information \mathcal{S}_t generated by current and past signals. For the following three reasons, it is interesting to construct an implied g_{t+1}^* from the m_{t+1}^* associated with R_t^1 or T^1 and the h_t^* associated with R_t^2 or T^2 .

First, actions depend only on signal histories. Hidden states are used either to depict the underlying uncertainty or to help represent preferences. However, agents cannot take actions contingent on these hidden states, only on the signal histories.

Second, in decentralized economies, asset prices can be characterized by stochastic discount factors that equal the intertemporal marginal rates of substitution of investors who are off corners and that depend on the distorted probabilities these investors use to value contingent claims. Since contingent claims to consumption can depend only on signal histories (and not on hidden states), the distortion to the signal distribution is the twist to asset pricing that is contributed by investors' concerns about model misspecification. In particular, under the approximating model, $\frac{g_{t+1}}{E[g_{t+1}|\mathcal{S}_t]}$ becomes a multiplicative adjustment to the ordinary stochastic discount factor for a representative agent (e.g., see [25]). It follows that the temporal inconsistency of worst case beliefs over hidden states discussed in Section 6.4 does not prevent appealing to standard results on the recursive structure of asset pricing in settings with complete markets.¹⁸

Third, Anderson et al. [2] found it useful to characterize detection probabilities using relative entropy and an alternative measure of entropy due to [8]. Chernoff [8] showed how detection error probabilities for competing models give a way to measure model discrepancy. Models are close when they are hard to distinguish with historical data. Because signal histories contain all data that are available to a decision maker, the measured entropy from distorting the signal distribution is pertinent for statistical discrimination. These lead us to measure either $E(g_{t+1}^* \log g_{t+1}^* | \mathcal{S}_t)$ or Chernoff's counterpart, as in [2].¹⁹

¹⁸ See Johnson and Donaldson [29].

¹⁹ Anderson et al. [2] show a close connection between the market price of risk and a bound on the error probability for a statistical test for discriminating the approximating model from the worst case model.

Our characterizations of worst case models have conditioned implicitly on the current period action. The implied distortion in the signal density is

$$\int \phi_t(z^*, s^*) \tau(z^*, s^* | y_t, z, a_t) \psi_t(z) q_t(z) d\lambda(z^*) d\lambda(z),$$

where ϕ_t is given by formula (13) and ψ_t is given by (16). When a Bellman–Isaacs condition is satisfied,²⁰ we can substitute for the control law and construct a conditional worst case conditional probability density for s_{t+1} as a function of the Markov state (y_t, q_t) . The process $\{(y_{t+1}, q_{t+1}) : t \geq 0\}$ is Markov under the worst case distribution for the signal evolution. The density q_t remains a component of the state vector.

8. A recursive multiple priors model

To attain a notion of dynamic consistency when the decision maker has multiple models, Epstein and Schneider [16,17] advocate a formulation that, when translated into our setting, implies time varying values for θ_1 and θ_2 . Epstein and Schneider advocate sequential constraints on sets of transition probabilities for signal distributions. To implement their proposal in our context, we can replace our fixed penalty parameters θ_1, θ_2 with two sequences of constraints on relative entropy.

In particular, suppose that

$$\varepsilon_t^1(m_{t+1}) \leq \kappa_t^1, \tag{31}$$

where κ_t^1 is a positive random variable in \mathcal{X}_t , and

$$\varepsilon_t^2(h_t) \leq \kappa_t^2, \tag{32}$$

where κ_t^2 is a positive random variable in \mathcal{S}_t . If these constraints bind, the worst-case probability distributions are again exponentially tilted. We can take θ_t^1 to be the \mathcal{X}_t -measurable Lagrange Multiplier on constraint (31), where $m_{t+1}^* \propto \exp\left(-\frac{W_{t+1}}{\theta_t^1}\right)$ and θ_t^1 solves $\varepsilon_t^1(m_{t+1}^*) = \kappa_t^1$. The counterpart to $R_t^1(W_{t+1})$ is

$$C_t^1(W_{t+1}) \doteq \frac{E \left[W_{t+1} \exp \left(-\frac{W_{t+1}}{\theta_t^1} \right) \middle| \mathcal{X}_t \right]}{E \left[\exp \left(-\frac{W_{t+1}}{\theta_t^1} \right) \middle| \mathcal{X}_t \right]}.$$

Similarly, let θ_t^2 be the \mathcal{S}_t -measurable Lagrange multiplier on constraint (32), where $h_t^* \propto \exp\left(-\frac{\hat{W}_t}{\theta_t^2}\right)$, and θ_t^2 solves $\varepsilon_t^2(h_t^*) = \kappa_t^2$. The counterpart to $R_t^2(\hat{W}_t)$ is

$$C_t^2(\hat{W}_t) \doteq \frac{E \left[\hat{W}_t \exp \left(-\frac{\hat{W}_t}{\theta_t^2} \right) \middle| \mathcal{S}_t \right]}{E \left[\exp \left(-\frac{\hat{W}_t}{\theta_t^2} \right) \middle| \mathcal{S}_t \right]}.$$

These constraint problems lead to natural counterparts to the operators T^1 and T^2 .

²⁰ For example, see Hansen et al. [26] or Hansen and Sargent [24].

Constraint formulations provide a justification for making θ_1 and θ_2 state- or time-dependent. Values of θ_1 and θ_2 would coincide if the two constraints were replaced by a single entropy constraint $E [h_t \varepsilon_t^1(m_{t+1}) | \mathcal{S}_t] + \varepsilon_t^2(h_t) \leq \kappa_t$, where κ_t is \mathcal{S}_t -measurable. Lin et al. [38] and Maenhout [39] give other reasons for making the robustness penalty parameters state dependent.²¹ With such state dependence, it can still be useful to disentangle misspecifications of the state dynamics and the distribution of the hidden state given current information. Using separate values for θ_1 and θ_2 achieves that.

9. Risk sensitivity and compound lotteries

Jacobson [28] linked a concern about robustness, as represented in the first line of (10), to risk sensitivity, as conveyed in the second line of (10). That link has been exploited in the control theory literature, for example, by Whittle [44]. Our desire to separate the concern about misspecified state dynamics from concern about misspecifying the distribution of the state inspires two risk-sensitivity operators. Although our primary interest is to let a decision maker respond to model misspecification, our two operators can also be interpreted in terms of enhanced risk aversion.²²

9.1. Risk-sensitive interpretation of R_t^1

The R_t^1 operator has an alternative interpretation as a *risk-sensitive* adjustment to continuation values that expresses how a decision maker who has no concern about robustness prefers to adjust continuation values for their risk. The literature on risk-sensitive control uses adjustments of the same $\log E \exp$ form that emerge from an entropy penalty and a concern for robustness, as asserted in (10). There are risk adjustments that are more general than those of the $\log E \exp$ form associated with risk-sensitivity. In particular, we could follow Kreps and Porteus [36] and Epstein and Zin [18] in relaxing the assumption that a temporal compound lottery can be reduced to a simple lottery without regard to how the uncertainty is resolved, which would lead us to adjust continuation values by

$$\tilde{R}_t^1(V_{t+1}) = \phi^{-1} (E [\phi(V_{t+1}) | \mathcal{X}_t])$$

for some concave increasing function ϕ . The risk-sensitive case is the special one in which ϕ is an exponential function. We focus on the special risk-sensitivity $\log E \exp$ adjustment because it allows us to use entropy to interpret the resulting adjustment as a way of inducing robust decision rules.

9.2. R_t^2 and the reduction of compound lotteries

While (16) shows that the operator R_t^2 assigns a worst-case probability distribution, another interpretation along the lines of Segal [41], Klibanoff et al. [35], and Ergin and Gul [19] is available. This operator adjusts for state risk differently than does the usual Bayesian model averaging

²¹ These authors consider problems without hidden states, but their motivation for state dependence would carry over to decision problems with hidden states.

²² Using detection probabilities, Anderson et al. [2] describe alternative senses in which the risk-sensitivity and robustness interpretations are and are not observationally equivalent. We intend eventually to study the interesting issues that arise in extending detection error probabilities to discipline the choice of θ_1, θ_2 pairs.

approach. Specifically, we can regard the transformation R_t^2 as a version of what Klibanoff et al. [35] call constant ambiguity aversion. More generally, we could use

$$\tilde{R}_t^2(\hat{V}_t) = \psi^{-1} E \left[\psi(\hat{V}_t) | \mathcal{S}_t \right]$$

for some concave increasing function ψ . Again, we use the particular ‘log $E \exp$ ’ adjustment because of its explicit link to entropy-based robustness.

10. Another example

Hansen and Sargent [22] follow Tallarini [42] and start with a representative consumer who, if he did not want to make a risk-sensitivity or robustness adjustment, would value streams of log consumption c_t according to

$$V_t = (1 - \beta)c_t + E_t \beta V_{t+1}.$$

But he wants to make multiple risk sensitivity adjustments to reflect multiple doubts about the stochastic specification of consumption growth. The representative consumer has two submodels for consumption growth $s_{t+1} \equiv c_{t+1} - c_t$, each of which has the state space form

$$\begin{aligned} \zeta_{t+1}(i) &= A(i)\zeta_t(i) + C(i)w_{t+1}, \\ c_{t+1} - c_t \equiv s_{t+1} &= D(i)\zeta_t(i) + G(i)w_{t+1}, \end{aligned}$$

where $\{w_{t+1}\}$ is an iid Gaussian process with mean 0 and covariance I and $\zeta_0(i)$ is normally distributed with mean $\zeta_0(i)$ and covariance matrix $\Sigma_0(i)$. Denote the submodels $i \in \{0, 1\}$ and suppose that the representative consumer attaches probability $\check{p}_t = E(i | \mathcal{S}_t)$ to model 1 at time t . These probabilities can be computed by using Bayes rule and data $s^t = [s_t, s_{t-1}, \dots, s_1]$.

Hansen and Sargent [22] specify submodel $i = 0$ so that it makes consumption growth be an i.i.d. Gaussian process with an unknown mean. Submodel $i = 1$ is like, but not identical to, a model of Bansal and Yaron [3] that makes consumption growth contain a difficult to detect persistent component. In addition to the uncertainty about shocks w_{t+1} assumed by Bansal and Yaron, one component of $\zeta(1)$ is a constant conditional mean of consumption that is unknown to the representative consumer. This feature would increase the risk faced by our representative consumer relative to Bansal and Yaron’s, even if he set $\hat{p}_0 = 1$. The representative consumer learns about the mean consumption growth parameters as well as other parts of the hidden state $z_t = [\zeta_t(0), \zeta_t(1), i]$.

The results of applying Bayes’ law to submodel i can be represented in terms of an innovations representation that takes the form

$$\begin{aligned} \check{\zeta}_{t+1}(i) &= A(i)\check{\zeta}_t(i) + K[\Sigma_t(i), i]\check{w}_{t+1}(i), \\ \Sigma_{t+1}(i) &= A(i)\Sigma_t(i)A(i)' + C(i)C(i)' - K[\Sigma_t(i), i][A(i)\Sigma_t(i)A(i)' + C(i)G(i)G(i)'], \\ s_{t+1} &= D(i)\check{\zeta}_t(i) + \check{w}_{t+1}(i), \end{aligned}$$

where

$$K[\Sigma_t(i), i] \doteq [A(i)\Sigma_t(i)D(i)' + C(i)G(i)G(i)'] [D(i)\Sigma_t(i)D(i)' + G(i)G(i)']^{-1},$$

$\check{\zeta}_{t+1}(i) = E[\zeta_{t+1} | s^t, i]$, $\check{w}_{t+1}(i)$ is the forecast error for the signal (i.e., the ‘innovation’), and $\Sigma_t(i)$ is the covariance matrix for $\zeta_t(i) - \check{\zeta}_t(i)$ conditioned on i and the signal history through

date t . Evidently, in this model, $\check{\zeta}_t(i)$, $\Sigma_t(i)$, $\iota = 0, 1$, and \check{p}_t are sufficient statistics for the joint distribution $q_t(z)$.

Hansen and Sargent [22] apply recursions (20), (21) to form the stochastic discount factor implied by a representative consumer who is concerned about misspecifications of the following distributions: (i) the distributions of (z_{t+1}, s_{t+1}) conditioned on $[t, \zeta_t(i)]$; (ii) the distributions of $\zeta_t(i)$ conditioned on $[t, \mathcal{S}_t]$; and (iii) the distributions of ι , conditional on \mathcal{S}_t . The representative consumer of Hansen and Sargent [22] applies T^1 to adjust for his suspicion about (i) and iterates on (20) to find valuations as functions of $\check{\zeta}(i)$, ι . The representative consumer makes adjustment (21) by applying T^2 first to adjust the distribution mentioned in (ii). He then applies another T^2 operator to adjust for suspicion of the distribution mentioned in (iii).²³ The implied Radon–Nikodym derivative that perturbs the distribution of $s_{t+1} = c_{t+1} - c_t$ conditional on \mathcal{S}_t serves as a multiplicative adjustment to the stochastic discount factor; in a T^1 -only model, Hansen et al. [25] dubbed its conditional standard deviation the market price of model uncertainty. Hansen and Sargent [22] study market prices of model uncertainty that emerge from the setting described here and investigate how it compares to ones that emerge from the T^1 only models of Hansen et al. [25] and Tallarini [42].

The distributions mentioned in (i) and (ii) of the previous paragraph are both Gaussian, while the one in (iii) is a scalar $\in (0, 1)$. The logarithmic preference specification makes the value function for the problem posed in Section 5.1.1 be affine in $\check{\zeta}$, c . As a result, the calculations in this model become very easy—the Kalman filter does the hard work in implementing Bayes' Law and the calculations of T^1 , T^2 for the linear-quadratic Gaussian model in Hansen and Sargent [23] apply. The assumption that $A(i)$, $C(i)$, $D(i)$, $G(i)$ are known accounts for this simplicity. Extending the model to let some elements in these matrices be unknown enriches the scope for modeling learning about unknown parameters at the cost of making the filtering problem nonlinear and so pushing it beyond the range of the Kalman filter. Hansen et al. [27] study such problems.

11. Concluding remarks

By incorporating learning, this paper responds to thoughtful criticisms of our earlier work about recursive formulations of robust control without learning. The framework here allows us to examine the consequences for valuations and decision rules of learnable components of the state that can capture both model selection and parameter estimation.

The model in Section 10 is about a pure endowment economy, so that the representative consumer chooses no actions—his worst case model determines valuations but not actions. Of course, the framework in this paper allows us also to study settings in which a decision maker chooses an action that influences the motion of the state. We illustrate this aspect by performing an analysis of robust experimentation in [12]. For a given concern about misspecification of hidden state probabilities as measured by θ_2 , we can study the speed at which learning diminishes concerns about misspecification along particular dimensions of uncertainty as the accretion of data together with Bayes' law gradually reduces the set of perturbed models by tightening posterior probabilities. The formulas in [23,22] show precisely how the volatilities of hidden state estimates that come from Bayes' law affect the gap between the worst case probabilities and those from the approximating model.

²³ Notice that by using different θ_2 's in these two applications of T^2 , we could focus the decision maker's concerns about robustness more on one of the two potential sources of misspecification.

Our procedures for solving robust discounted dynamic programming problems are as easy to use as corresponding problems without concerns about robustness and come down to replacing each of two conditional expectations operators in the problem without robustness with a risk-sensitivity operator. For a finite θ_1 , the operator T^1 captures the decision maker's fear that the state and signal dynamics conditioned on both the observed and hidden components of the state are misspecified. For a finite θ_2 , the operator T^2 captures the decision maker's fear that the distribution of the hidden state conditioned on the history of signals is misspecified. Using different values of θ_1 and θ_2 in the operators T^1 and T^2 gives us the freedom to focus distrust on different aspects of the decision maker's model.²⁴

We do not address the interesting issues that would arise in an economy with heterogeneous agents who have different specification concerns about the same approximating model. Anderson [1] studies how Pareto optimal allocations for such economies put history dependence into Pareto weights. Anderson does not ascribe learning problems to his agents, but it would be interesting to study them in such heterogeneous agent contexts.

Acknowledgments

We thank In-Koo Cho, Anastasios Karantounias, Kenneth Kasa, Ricardo Mayer, Grace Tsiang, Tomasz Piskorski, the editor, and especially a referee for helpful comments on earlier drafts of this paper. This material is based on work supported by the National Science Foundation under Award Numbers SES0519372 and SES0137035.

References

- [1] E.W. Anderson, The dynamics of risk sensitive allocations, *J. Econ. Theory* 125 (2005) 93–150.
- [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. Econ. Assoc.* 1 (1) (2003) 68–123.
- [3] R. Bansal, A. Yaron, Risks for the long run: a potential resolution of asset pricing puzzles, *J. Finance* LIX (4) (2004) 1481–1509.
- [4] T. Başar, P. Bernhard, *H[∞]-Optimal Control and Related Minimax Design Problems*, second ed., Birkhauser, Basel, 1995.
- [5] D. Bergemann, J. Valimaki, Learning and Strategic Pricing, *Econometrica* 64 (1996) 1125–1149.
- [6] W.A. Brock, S.N. Durlauf, K.D. West, Policy evaluation in uncertain economic environments, 2003.
- [7] W.A. Brock, S. Durlauf, K.D. West, Model uncertainty and policy evaluation: some theory and empirics, University of Wisconsin, SSRI paper number 2004-19, 2004.
- [8] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Statist.* 23 (1952) 493–507.
- [9] I.-K. Cho, K. Kasa, Model validation dynamics and endogenous currency crises, Unpublished, 2005.
- [10] I.-K. Cho, K. Kasa, Learning and model validation, Unpublished, 2006.
- [11] I.-K. Cho, N. Williams, T.J. Sargent, Escaping nash inflation, *Rev. Econ. Stud.* 69 (2002) 1–40.

²⁴ Specifications with $\theta_1 = \theta_2$ emerge when we follow Hansen and Sargent [21] by adopting a timing protocol that requires the malevolent agent to commit to a worst case model $\{M_{t+1}\}$ once and for all at time 0. Hansen and Sargent [21] give a recursive representation for the solution of the commitment problem in terms of R_t^1 and R_t^2 operators with a common but time-varying multiplier equal to $\frac{\theta}{\beta^t}$. The presence of β^t causes the decision maker's concerns about misspecification to vanish for tail events. Only for the undiscounted case does the zero-sum two player game with commitment in [21] give identical outcomes to the games without commitment in this paper. As noted in Section 6.6, when $\beta < 1$, the gap between the outcomes with and without commitment is the source of time-inconsistency of the worst case beliefs about the hidden state. Much of the control theory literature (e.g. [44,4]) uses the commitment timing protocol and sets $\beta = 1$. Hansen and Sargent [21] show how to represent parts of that literature in terms of our formulation of model perturbations as martingales.

- [12] T. Cogley, R. Colacito, L. Hansen, T. Sargent. Robustness and U.S. monetary policy experimentation, Unpublished, 2005.
- [13] T. Cogley, R. Colacito, T. Sargent, Benefits from U.S. monetary policy experimentation in the days of Samuelson and Solow and Lucas, *J. Money Credit Bank.* 39 (Suppl. 1) 67–99.
- [14] P. Dupuis, R.S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley Series in Probability and Statistics, Wiley, New York, 1997.
- [15] R.J. Elliott, L. Aggoun, J.B. Moore, *Hidden Markov Models: Estimation and Control*, Springer, New York, 1995.
- [16] L. Epstein, M. Schneider, Independently and indistinguishably distributed, *J. Econ. Theory* 113 (1) (2003) 32–50.
- [17] L. Epstein, M. Schneider, Recursive multiple priors, *J. Econ. Theory* 113 (1) (2003) 1–31.
- [18] L. Epstein, S. Zin, Substitution risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework, *Econometrica* 57 (1989) 937–969.
- [19] H. Ergin, F. Gul, A subjective theory of compound lotteries, Unpublished, 2004.
- [20] L.P. Hansen, T. Sargent, Discounted linear exponential quadratic Gaussian control, *IEEE Trans. Automat. Control* 40 (5) (1995) 968–971.
- [21] L.P. Hansen, T.J. Sargent, Robust estimation and control under commitment, *J. Econ. Theory* 124 (2) (2005) 258–301.
- [22] L.P. Hansen, T.J. Sargent, Fragile beliefs and the price of model uncertainty, Unpublished, 2006.
- [23] L.P. Hansen, T.J. Sargent, Robust estimation and control for LQ Gaussian problems without commitment, Unpublished, 2006.
- [24] L.P. Hansen, T.J. Sargent, *Robustness*, Princeton University Press, 2006, Princeton, NJ, forthcoming.
- [25] L.P. Hansen, T. Sargent, T. Tallarini, Robust permanent income and pricing, *Rev. Econ. Stud.* 66 (1999) 873–907.
- [26] 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.
- [27] L.P. Hansen, N. Polson, T.J. Sargent, Fragile Beliefs with parameter estimation, Unpublished, 2006.
- [28] D.H. Jacobson, Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games, *IEEE Trans. Automat. Control* AC-18 (1973) 1124–1131.
- [29] T.H. Johnsen, J.B. Donaldson, The structure of intertemporal preferences under uncertainty and time consistent plans, *Econometrica* 53 (1985) 1451–1458.
- [30] B. Jovanovic, Job matching and the theory of turnover, *J. Polit. Econ.* 87 (5) (1979) 972–990.
- [31] B. Jovanovic, Selection and the evolution of industry, *Econometrica* 50 (3) (1982) 649–670.
- [32] B. Jovanovic, Y. Nyarko, The transfer of human capital, *J. Econ. Dynam. Control* 19 (1995) 1033–1064.
- [33] B. Jovanovic, Y. Nyarko, Learning by doing and the choice of technology, *Econometrica* 64 (6) (1996) 1299–1310.
- [34] K. Kasa, Learning, large deviations, and recurrent currency crises, *Int. Econ. Rev.* 45 (2004) 141–173.
- [35] P. Klibanoff, M. Marinacci, S. Mukerji, *A smooth model of decision making under ambiguity*, Northwestern University, 2003.
- [36] D.M. Kreps, E.L. Porteus, Temporal resolution of uncertainty and dynamic choice, *Econometrica* 46 (1978) 185–200.
- [37] D. Laibson, Golden eggs and hyperbolic discounting, *Quart. J. Econ.* 112 (2) (1997) 443–478.
- [38] J. Lin, J. Pan, T. Wang, An equilibrium model for rare-event premia and its implication for option pricing, *Rev. Finan. Stud.* 18 (1) (2005) 131–164.
- [39] P.J. Maenhout, Robust portfolio rules and asset pricing, *Rev. Finan. Stud.* 17 (4) (2004) 951–983.
- [40] I.R. Petersen, M.R. James, P. Dupuis, Minimax optimal control of stochastic uncertain systems with relative entropy constraints, *IEEE Trans. Automat. Control* 45 (2000) 398–412.
- [41] U. Segal, Two-stage lotteries without the reduction axiom, *Econometrica* 58 (1990) 349–377.
- [42] T.D. Tallarini, Risk-sensitive real business cycles, *J. Monetary Econ.* 45 (3) (2000) 507–532.
- [43] V. Weiland, Comment on certainty equivalence and model uncertainty, in: A. Orphanides, D. Reifschneider, J. Faust (Eds.), *Models and Monetary Policy Research in the Tradition of Dale Henderson, Richard Porter and Peter Tinsley*, Board of Governors of the Federal Reserve System, Washington, DC, 2005.
- [44] P. Whittle, *Risk-sensitive Optimal Control*, Wiley, New York, 1990.
- [45] W.J. Wonham, Some applications of stochastic differential equations to optimal nonlinear filtering, *SIAM J. Control* 2 (1964) 347–368.