

Robust Estimation and Control under Commitment

Lars Peter Hansen*
Thomas J. Sargent[‡]

March 10, 2005

Abstract

In a Markov decision problem with hidden state variables, a decision maker expresses fear that his model is misspecified by surrounding it with a set of alternatives that are nearby as measured by their expected log likelihood ratios (entropies). Sets of martingales represent alternative models. Within a two-player zero-sum game under commitment, a single minimizing player chooses a martingale at time 0. Probability distributions that solve distorted filtering problems serve as state variables, much like the posterior in problems without concerns about misspecification. We display conditions under which an equilibrium of the zero-sum game with commitment has a recursive representation, then cast that representation in terms of two risk-sensitivity operators. We apply our results to a linear quadratic example that makes contact with the analysis of Basar and Bernhard (1995) and Whittle (1990).

1 Introduction

In single agent problems with incomplete information, optimal decision rules depend on a decision maker's posterior distribution over hidden state variables, called $q_t(z)$ here, an object that summarizes the pertinent history of observed signals. The decision maker expresses faith in his model when he uses Bayes' rule to deduce the transition law for $q_t(z)$.¹

But how should a decision maker proceed if he doubts his model and wants a decision rule that is robust to a set of statistically difficult to detect misspecifications of it? This paper studies a decision maker who makes his fear of model misspecification concrete by surrounding his approximating model with the set of all alternative models whose expected log likelihood ratios (i.e., whose *relative entropies*) are restricted or penalized. If the relative entropies are constrained to be small, the decision maker believes that his model is a good

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

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

[‡]We thank Tomasz Piskorski, Gahuar Turmuhambetova and Fernando Alvarez for especially helpful comments on earlier drafts of this paper.

¹For example, see Jovanovic (1979), Jovanovic (1982), Jovanovic and Nyarko (1995), Jovanovic and Nyarko (1996), Bergemann and Valimaki (1996), Elliott, Aggoun, and Moore (1995), and Cogley, Colacito, and Sargent (2005).

approximation. The decision maker wants robustness against these alternatives because, as Anderson, Hansen, and Sargent (2003) emphasize, perturbations with small relative entropies are statistically difficult to distinguish from the approximating model. This paper provides an appropriate lower dimensional object to summarize the history of signals when the decision maker wants a decision rule to perform well for this set of models. We study how the appropriate summary of the history of signals depends on details of the decision maker's problem, including how he discounts future utilities and contributions to entropy, whether hidden states enter the decision maker's period utility function, and whether the decision maker chooses sequentially or once and for all at time 0. We describe special circumstances under which the appropriate summary of signals continues to be the decision maker's posterior under the approximating model, despite the fact that he distrusts his model.

All formulations of robust decision making in dynamic contexts begin with a zero sum game in which a minimizing player chooses a worst-case model from a set surrounding the approximating model. Alternative formulations employ different timing protocols and different ways of specifying the set of alternative models. This paper adopts a timing protocol that prevails throughout much of the literature on robust control. It ascribes commitment to both players in the zero sum game. To create a counterpart to the recursive formulation of a zero-sum two-player game that extends the formulation in Hansen and Sargent (1995) to situations with hidden states, Hansen and Sargent (2005) analyzes games without commitment.

The remainder of this paper is organized as follows. 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 can be difficult to distinguish from it statistically, and construct a decision rule that can work well for all of those models. Section 3 describes how to represent distortions of an approximating model in terms of martingales defined on the same probability space as the approximating model. This section then defines two operators, R_1 and R_2 , that are indexed by a common penalty parameter $\theta \in (\underline{\theta}, +\infty)$ and that adjust expectations of continuation values, viewed as random variables measurable with respect to histories of states and signals, respectively, for lack of confidence in the approximating model. We interpret θ as a penalty on an entropy term that measures the size of allowable model misspecifications. Subsequent sections use those two operators, or closely related ones, to construct decision rules that are robust to departures from the approximating model. Section 4 expresses a robust control problem in terms of our section 3 concept of discounted entropy and formulates a robust control problem under commitment, meaning that the worst-case model and robust decision rule are both chosen at time 0 and are never reconsidered. Section 5 briefly discusses the connection of the section 4 setup to work by Chamberlain (2000) and Knox (2003).

To highlight time consistency issues and to set the stage for work in Hansen and Sargent (2005) that define alternative zero-sum games recursively, the formulation in section 4 allows different factors β for discounting one-period utilities and ρ for discounting one-period contributions to entropy. As a prolegomenon to more general formulations without commitment, section 6 studies the special case in which $\rho = 1$. When $\rho = 1$, we can implement the solution of the commitment problem recursively through appropriate iterated applications of R_1 and R_2 . We thereby discover an appropriate counterpart to the distribution over the

hidden state variables $q_t(z)$ that occurs in the basic problem in section 2 without a concern for robustness.

Section 6 shows that when $\rho = 1$, concerns about robustness wear off, a feature that we sometimes want to avoid. Hansen and Sargent (2005) show how setting $\rho = \beta$ sustains enduring concerns about robustness. But they must confront the impediment that when $\rho \neq 1$, our section 4 results giving a recursive representation of a commitment solution do not apply. To make concerns about robustness to endure, Hansen and Sargent (2005) have to accept a form of time-inconsistency beliefs about hidden states (but not about signals).

Section 7 applies our results to the classic linear-quadratic stochastic robust control problem. For the case $\rho = 1$, we describe how to implement the solution of the commitment problem recursively, describe an object $\check{q}_t(z)$ that emerges from the application of the \mathbf{R}^2 operator, and link $\check{q}_t(z)$ to the distribution over hidden states conditioned on the history of signals, namely, $q_t(z)$, that emerges from applying the ordinary Kalman filter. In an important special case in which hidden states do not appear in the one-period return function, $\check{q}_t(z) = q_t(z)$.

Section 8 concludes by mentioning how Hansen and Sargent (2005) create a recursive model of robust decision making by using different multiplier parameters θ_1 and θ_2 to define Markov versions of the two operators \mathbf{R}^1 and \mathbf{R}^2 . By using different θ 's, they focus the decision maker's fear of misspecification more on either the distribution of the hidden state vector or on the transition dynamics themselves. This extension is especially useful in continuous time formulations. Appendix A confirms assertions of useful properties of martingales. Appendix B verifies the link between the solution to the linear quadratic commitment problem as we have posed it with a solution from the control theory literature.

1.1 Related literature

In a class of dynamic robust parameter estimation problems, Chamberlain (2000) and Knox (2003) solve what we interpret as a date zero commitment game with a time-invariant hidden state. Because they want to capture the idea that the decision maker cannot commit to a robust choice of a prior distribution at date zero, Epstein and Schneider (2003) advocate a timing protocol for a zero-sum game that differs from Chamberlain's and Knox's: Epstein and Schneider formulate the estimation problem sequentially.

Epstein and Schneider (2003) and Knox (2003) mention axioms. We do not start from axioms but instead purposefully design our procedures to make contact with the literatures on robust control and estimation.² Thus, we begin by studying a commitment problem that is close to one from the literature on robust control. We adopt this starting point partly because the robust control literature contains so many useful conceptual and computational insights. We use computational tricks from that literature to get a recursive formulation that is computationally tractable. Like Chamberlain's and Knox's, the problem solved in the control theory literature is a date zero commitment problem. But in an interesting special case, the solution has a convenient recursive representation that we use to highlight the roles of commitment and discounting in decision problems with hidden state variables. We then alter the assumptions about discounting and commitment. Although we also believe that

²Maccheroni, Marinacci, and Rustichini (2004) have developed axioms that justify static versions of some of our procedures.

solutions to the commitment problem are interesting, we tell why we like the no-commitment assumption and some properties of its consequences.

Much of the literature on robust control poses problems in a non-stochastic context. (See Petersen, James, and Dupuis (2000) for one of the initial stochastic formulations.) Following Hansen, Sargent, Turmuhambetova, and Williams (2004), we pose a stochastic version of a robust control problem by using a martingale formulation to distort the probability distribution associated with an approximating model. In Hansen and Sargent (2005), we discuss how our formulation relates to ones advocated by Epstein and Schneider (2003).

2 The approximating model

This section formulates the Bellman equation for a standard recursive decision problem. The presence of hidden state variables impels us to include a law of motion for the distribution of hidden state variables. An approximating model includes the motion of the posterior distribution over hidden states. Subsequent sections acknowledge and decompose doubts about the approximating model.

A state vector $\{x_t : t \geq 0\}$ is Markov with a time invariant transition density that can be influenced by a control process $\{a_t : t \geq 0\}$. Partition the state vector as

$$x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}$$

where y_t is observed and z_t is not; y_t is a component of a possibly larger vector s_t of signals that inform the decision maker about the unobserved components z_t . 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 generated by y_0 and current and past signals s , and let $\{\mathcal{X}_t : t \geq 0\}$ be a larger filtration that includes information generated by the history of x . The smallest sigma algebra generated by all states including the future is:

$$\mathcal{X}_\infty \doteq \bigvee_{t \geq 0} \mathcal{X}_t,$$

and similarly the smallest sigma algebra generated by all signals 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. Let \mathcal{A}_t^* be the set of A -valued random vectors that are \mathcal{X}_t measurable.

We assume that the conditional joint probability distribution for states and signals has the following structure:

- A law of motion for the evolution of the observable states:

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

When y_0 is known, this law allows us to construct an observable state recursively from signals and actions.

- Two components that determine the evolution of hidden states and signals:

(a) An exogenously specified density $\tau(z^*, s^* | x_t, a_t)$; τ is a density relative to the product measure $\lambda \times \eta$. For notational simplicity, we will assume that τ is the conditional density implied by:

$$\begin{aligned} s_{t+1} &= \pi_s(x_t, a_t, w_{t+1}) \\ z_{t+1} &= \pi_z(x_t, a_t, w_{t+1}) \end{aligned}$$

where $\{w_{t+1} : t \geq 0\}$ is an i.i.d. sequence of random variables. Via Bayes' rule, τ implies the second component, namely,

(b) A density q_t for z_t conditioned on information \mathcal{S}_t generated by y_0, s_1, \dots, s_t ; q_t is a density relative to the measure λ .

We are interested in the following decision problem under incomplete information about the state:

Problem 2.1.

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

subject to

$$\begin{aligned} y_{t+1} &= \pi_y(s_{t+1}, y_t, a_t) \\ z_{t+1} &= \pi_z(x_t, a_t, w_{t+1}) \\ s_{t+1} &= \pi_s(x_t, a_t, w_{t+1}). \end{aligned}$$

It is possible to express problem 2.1 in a more convenient way by focusing separately on the control and estimation aspects of the problem. To prepare the way, we introduce the following structure.

We use τ to construct two densities for the signal, the first of which is conditioned on a finer information set:

$$\begin{aligned} \kappa(s^* | y_t, z_t, a_t) &\doteq \int \tau(z^*, s^* | y_t, z_t, a_t) d\lambda(z^*) \\ \varsigma(s^* | y_t, q_t, a_t) &\doteq \int \kappa(s^* | y_t, z, a_t) q_t(z) d\lambda(z). \end{aligned} \tag{1}$$

The decomposition on the right side of (1) will play an important role. By Bayes' rule:

$$\begin{aligned} q_{t+1}(z^*) &= \frac{\int \tau(z^*, s_{t+1} | y_t, z, a_t) q_t(z) d\lambda(z)}{\varsigma(s_{t+1} | y_t, q_t, a_t)} \\ &\equiv \pi_q(s_{t+1}, y_t, q_t, a_t). \end{aligned}$$

In particular applications, the mapping π_q can be computed by using filtering methods that specialize Bayes' rule (e.g., the Kalman filter or the Wonham filter).

In defining the *state* of a recursive decision problem, the information structure and Bayes' rule direct us to replace the unobservable z_t with its density q_t conditional on current and past signals. We therefore consider recursive formulations of a control problem with state (y_t, q_t) and transition law

$$y_{t+1} = \pi_y(s_{t+1}, y_t, a_t) \quad (2)$$

$$q_{t+1} = \pi_q(s_{t+1}, y_t, q_t, a_t). \quad (3)$$

Let

$$\pi = \begin{bmatrix} \pi_y \\ \pi_q \end{bmatrix}.$$

By using this structure, we can rewrite problem 2.1 in the alternative form:

Problem 2.2. Choose a sequence of decision rules for a_t as a function 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 a given density $q_0(z)$, the conditional density $\kappa(s_{t+1}|y_t, z_t, a_t)$, and equations (2)-(3). The Bellman equation for this problem is

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

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

Example 2.3. We modify a stochastic growth model due to Brock and Mirman (1972) and Merton (1975) by including a hidden growth state. The technology is:

$$K_{t+1} = (S_t)^{1-\alpha} (K_t)^\alpha - C_t + (1 - \delta)K_t \quad (4)$$

where K_t is the date t capital stock, C_t is date t consumption, δ is the rate of depreciation, and α is a productivity parameter. The labor supply is fixed and S_t is a labor augmenting technology process.

We let s_t denote the growth rate of the technology shock process $\log S_t - \log S_{t-1}$ and suppose:

$$s_{t+1} = \zeta \cdot z_t + \sigma w_{t+1}^1$$

where w_{t+1}^1 is a standard normal random variable independent of z_t . Realizations of s_t are noisy indicators of alternative growth rate regimes. The n -dimensional vector ζ contains alternative conditional means for growth rates and z_t is a random vector conformable to ζ whose realizations are unit vectors e_i . The evolution of the n dimensional unit vectors e_i is described by a probability matrix $T_{ij} = \text{Prob}(z_{t+1} = e_j | z_t = e_i)$.

This model of the technology shock is a special case of the regime shift models of Sclove (1983) and Hamilton (1989) and is also a discrete-time counterpart to the nonlinear filtering model of Wonham (1964). The decision maker observes y_t but not z_t at date t .

The random growth in technology makes it convenient to scale variables by the level of the technology. Thus we rewrite the production function (4) as:

$$k_{t+1} = \exp(s_t - s_{t+1}) [(k_t)^\alpha - c_t + (1 - \delta)k_t]$$

where $k_t = K_t/S_t$ and $c_t = C_t/S_t$.

Preferences are represented by discounting a utility function U

$$U(C) = \frac{(C)^{1-\gamma}}{1-\gamma} = \frac{\exp[(1-\gamma)(\log c + s)]}{1-\gamma}$$

with a discount factor β .

To map the planning problem associated with this growth model into our general decision problem, we take $y_t = k_t$, s_t to be the signal, and the hidden state to be the growth rate regime ζ for the technology shock. The control variable is $a_t = c_t$.

3 Representing distorted distributions

This section describes how to represent distorted probability distributions by expressing them as perturbations of the decision maker's approximating model. We construct a discrete-time version of a martingale representation of distortions that we have used in continuous time (see Hansen, Sargent, Turmuhambetova, and Williams (2004)). Then we present recursive representations of those martingales and alternative recursive formulations of discounted entropy to prepare a recursive formulation of the decision problem with hidden state variables. We use these measures of entropy to induce two distorted expectations operators that are useful for constructing robust decision rules when some state variables are unobserved.

3.1 Simplified information structure

We will suppose that the decision maker cannot influence \mathcal{X}_t and that we can specify it exogenously. We allow the decision maker's actions to influence \mathcal{S}_t , though for convenience, we leave this dependence implicit in our notation. However, in problems in which decision maker's action doesn't influence the hidden state, either directly or indirectly, we can specify the filtration $\{\mathcal{S}_t : t \geq 0\}$ exogenously. Examples include problems in which a parameter or hidden state must be estimated.

3.2 Expectations and martingales

We represent a distorted probability measure that is absolutely continuous with respect to a baseline probability measure by using a Radon-Nikodym derivative. Consider a sample of length t and a distorted probability measure applied to \mathcal{X}_t . The Radon-Nikodym theorem implies the existence of a nonnegative \mathcal{X}_t measurable function M_t . To make it a probability measure, we impose $EM_t = 1$. Under the distorted probability distribution, the expectation of a bounded \mathcal{X}_t measurable random variable W_t is

$$\tilde{E}W_t \doteq EM_t W_t.$$

Recall that \mathcal{X}_{t+1} contains \mathcal{X}_t .

Definition 3.1. *The family of conditional expectation operators is consistent if the restriction of \tilde{E}_{t+1} to bounded, \mathcal{X}_t measurable random variables coincides with \tilde{E}_t for all $t \geq 0$.*

To attain consistency, we must have

$$EM_{t+1}W_t = EM_tW_t \quad (5)$$

for all bounded W_t that are \mathcal{X}_t measurable. This is equivalent to requiring that

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

We summarize this outcome in a well known result:

Lemma 3.2. *A family of distorted conditional expectation operators is consistent if and only if the process $\{M_t : t \geq 0\}$ is a martingale with unit expectation adapted to $\{\mathcal{X}_t : t \geq 0\}$.*

Remark 3.3. *Since the martingale $\{M_t : t \geq 0\}$ is nonnegative and has unit expectation for all t , Doob's martingale convergence theorem implies that it converges almost surely. Another notion of a limit comes from observing that our martingale formulation of the perturbed probabilities allows us to apply the Kolmogorov extension theorem. This establishes the existence of a probability distribution on \mathcal{X}_∞ that is consistent with those implied by the martingale on the sigma algebras $\{\mathcal{X}_t : t \geq 0\}$. Under the conditions imposed so far, these two limits need not be compatible. The probability measure implied by the Kolmogorov extension theorem does not have to be absolutely continuous with the respect to the probability distribution associated with the approximating model. Furthermore, the limiting random variable implied by the martingale convergence theorem does not necessarily have a unit expectation and hence is not necessarily a Radon-Nikodym for a probability measure. As we will see in section 3.4 and appendix A, these two limiting results can be linked when we establish a different form of martingale convergence.³*

3.3 Multiplicative decomposition of martingale

The martingale M_{t+1} allows us to represent distortions in non-sequential decision problems, i.e., problems under commitment. When we want to formulate sequential decision problems, it is convenient to decompose M_{t+1} into components that represent distortions to conditional probability distributions of date $t + 1$ events conditioned on date t information.⁴ 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

$$M_t = M_0 \prod_{j=1}^t m_j. \quad (6)$$

³Section 7 contains an example in which $\{M_t : t \geq 0\}$ has a limit that is the Radon-Nikodym derivative for the limiting probability measure. As a consequence the effects of a concern about robustness on decisions and estimates vanishes over time.

⁴This is like factoring a likelihood function.

The random variable M_0 has unconditional expectation equal to unity.

By construction, the random variable 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:

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

where the denominator is strictly positive. We use m_{t+1} to model 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 $E(m_{t+1}|\mathcal{X}_t) = 1$. In formulating robust control problems, we use (6) to construct a martingale from an appropriately restricted process $\{m_{t+1} : t \geq 0\}$ and an initial condition M_0 .

3.4 Entropy

The entropy of the distortion at time t conditioned on date zero information is $E(M_t \log M_t | \mathcal{X}_0)$. The product decomposition (6) of M_t implies a corresponding additive decomposition of entropy:

$$\begin{aligned} E(M_t \log M_t | \mathcal{X}_0) - E(M_0 \log M_0 | \mathcal{X}_0) &= \sum_{j=0}^{t-1} E(M_t \log m_{j+1} | \mathcal{X}_0) \\ &= \sum_{j=0}^{t-1} E[E(M_{j+1} \log m_{j+1} | \mathcal{X}_j) | \mathcal{X}_0] \\ &= \sum_{j=0}^{t-1} E[M_j E(m_{j+1} \log m_{j+1} | \mathcal{X}_j) | \mathcal{X}_0], \quad (7) \end{aligned}$$

where the second equality follows from (5) and the third from an application of the law of iterated expectations. We set $M_0 = 1$, which means that we explore only probability distortions conditioned on \mathcal{X}_0 .

Because $m \log m$ is convex in m ,

$$E(m_{t+1} \log m_{t+1} | \mathcal{X}_t) \geq E(m_{t+1} | \mathcal{X}_t) \log [E(m_{t+1} | \mathcal{X}_t)] = 0.$$

Definition 3.4. Consider a random variable m_{t+1} in \mathcal{M}_{t+1} . Its *conditional (on \mathcal{X}_t) relative entropy* is

$$\varepsilon_t^1(m_{t+1}) \doteq E(m_{t+1} \log m_{t+1} | \mathcal{X}_t). \quad (8)$$

Because all terms of (7) are nonnegative, the sequence

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

is increasing and has a limit that might be $+\infty$ with positive probability. Using this in (7) shows that

$$\lim_{t \rightarrow \infty} E(M_t \log M_t | \mathcal{X}_0)$$

converges. In the appendix, we show that when this limit is finite almost surely, the martingale sequence $\{M_t : t \geq 0\}$ converges in the sense that

$$\lim_{t \rightarrow \infty} E(|M_t - M_\infty| | \mathcal{X}_0) = 0, \quad (9)$$

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,

$$M_t = E(M_\infty | \mathcal{X}_t).$$

For short, we call the probability measure associated with the approximating model the ‘benchmark probability measure.’

Remark 3.5. *When M_t converges to a random variable M_∞ as in (9), the distorted probability distribution induced by M_∞ is absolutely continuous with respect to the benchmark probability distribution. Therefore, a Law of Large Numbers that obtains under the distorted model will also apply under the benchmark model.*

3.5 Discounted entropy

Where $0 < \rho \leq 1$, Hansen, Sargent, Turmuhambetova, and Williams (2004) use a continuous time analogue of $(1 - \rho) \sum_{t=1}^\infty \rho^t E(M_t \log M_t | \mathcal{X}_0)$ as a discounted measure of entropy for models without hidden states. To prepare recursive representations of some decision problems, it is convenient for us to use an alternative representation of discounted entropy inspired by the decomposition

$$\begin{aligned} (1 - \rho) \sum_{t=1}^\infty \rho^t E(M_t \log M_t | \mathcal{X}_0) &= \sum_{t=0}^\infty \rho^{t+1} E[M_t E(m_{t+1} \log m_{t+1} | \mathcal{X}_t) | \mathcal{X}_0], \\ &= \sum_{t=0}^\infty \rho^{t+1} E[M_t \varepsilon_t^1(m_{t+1}) | \mathcal{X}_0] \end{aligned} \quad (10)$$

where the right side is obtained by applying formulas (5), (6), and (7) and Abel’s summation formula (‘summation-by-parts’). The left side of (10) represents entropy in terms of the level distortions M_t , while the first term on the right side represents entropy in terms of the ratio distortions m_t . We can use contributions from both sides to form entropy penalties for robust control problems.⁵

When $E(M_t \log M_t | \mathcal{X}_0)$ converges, the limiting version of (10) as $\rho \rightarrow 1$ is

$$E(M_\infty \log M_\infty | \mathcal{X}_0) = \sum_{t=0}^\infty E[M_t E(m_{t+1} \log m_{t+1} | \mathcal{X}_t) | \mathcal{X}_0], \quad (11)$$

⁵The distinction between these level and difference measures of entropy will be especially revealing when we move to continuous time formulations.

where M_∞ is the limit point of the martingale. With discounting, the right side of (10) can be finite without convergence of the martingale sequence, and therefore without the common law of large numbers noted in remark 3.5.

3.6 Representation of induced distributions

To exploit the Markov structure of our recursive decision problem, it is useful to work with induced distributions of random vectors and to specify perturbations to those induced distributions directly. In this section, we briefly indicate how our approach of multiplying random variables by a positive random variable affects induced distributions. For concreteness, we consider the effect of multiplying by m_{t+1} before taking expectations over the induced conditional distribution for s_{t+1} .

Notice that the random variable m_{t+1} can be depicted as

$$\chi_t(z_{t+1}, s_{t+1}, x_t, \dots, x_0)$$

for some Borel measurable function χ_t and that

$$\int \chi_t(z, s, x_t, \dots, x_0) \tau(z, s | x_t) d\lambda(z) d\eta(s) = 1.$$

Therefore, associated with m_{t+1} is a multiplicative distortion in the density for (z_{t+1}, s_{t+1}) , given current information associated with the hidden and observed states. Formally, $\chi_t(\cdot | x_t, \dots, x_0)$ is a conditional density of a perturbed model with respect to the approximating model. The second of the following two equations expresses conditional relative entropy in terms of the distorted induced distribution:

$$\begin{aligned} \varepsilon_t^1(m_{t+1}) &\doteq E(m_{t+1} \log m_{t+1} | \mathcal{X}_t) \\ &= \int [\log \chi_t^*(z, s | x_t, x_{t-1}, \dots, x_0)] \chi_t^*(z, s | x_t, x_{t-1}, \dots, x_0) \tau(z, s | x_t) d\lambda(z) d\eta(s). \end{aligned}$$

3.7 Distorting likelihoods with hidden information

Because it is adapted to \mathcal{X}_t , the random variable M_t is a likelihood ratio for two probability distributions over \mathcal{X}_t . The implied likelihood ratio for the reduced information set \mathcal{S}_t is the \mathcal{S}_t -measurable random variable

$$G_t = E(M_t | \mathcal{S}_t)$$

that assigns distorted expectations to random variables in \mathcal{S}_t that agree with M_t ; $\{G_t : t \geq 0\}$ is a martingale adapted to the decision maker's information sequence $\{\mathcal{S}_t : t \geq 0\}$.

Conditioning on the smaller information set $\{\mathcal{S}_t : t \geq 0\}$ leads us to decompose M_t by first defining

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

The random variable h_t is \mathcal{X}_t measurable. The random variable M_t can be decomposed as

$$M_t = h_t G_t, \tag{12}$$

with a corresponding entropy decomposition:

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

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. In the decomposition (12), M_t distorts the probability distribution of \mathcal{X}_t , h_t distorts the probability of \mathcal{X}_t conditioned on \mathcal{S}_t , and G_t distorts the probability of \mathcal{S}_t .

Remark 3.6. *An entropy measure based on the larger information set \mathcal{X}_t exceeds that of the smaller information set \mathcal{S}_t , an inequality that is used in maximum likelihood estimation with hidden states and that underlies the EM algorithm.⁶*

Consider a distorted expectation operator that solves the following θ -penalized minimum entropy problem:

Problem 3.7.

$$\min_{M_t \geq 0, EM_t = 1,} E(M_t W_t | \mathcal{S}_0) + \theta E(M_t \log M_t | \mathcal{S}_0)$$

subject to the restriction that M_t be \mathcal{X}_t measurable.

In a decision problem in which actions depend on conditioning information in \mathcal{S}_t , it is convenient to use decomposition (12) and to solve for h_t before solving for G_t . Substituting (12) into the objective for problem 3.7 gives:

$$E(G_t h_t W_t | \mathcal{S}_0) + \theta E(G_t h_t \log h_t | \mathcal{S}_0) + \theta E(G_t \log G_t | \mathcal{S}_0).$$

As in other decision problems under uncertainty, h_t can be determined by solving the conditional problem:

Problem 3.8.

$$\min_{h_t \in \mathcal{H}_t} E(h_t W_t | \mathcal{S}_t) + \theta E(h_t \log h_t | \mathcal{S}_t),$$

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

After solving problem (3.8), we can return to problem (3.7) and compute the minimizing choice of G_t . In recursive formulations, when we want to focus only on actions contingent on \mathcal{S}_t and hence on the distortion h_t , we can skip the step of solving for G_t . This motivates our second measure of conditional relative entropy.

Definition 3.9. *Consider a random variable h_t in \mathcal{H}_t . Its **conditional (on \mathcal{S}_t) relative entropy** is*

$$\varepsilon_t^2(h_t) \doteq E(h_t \log h_t | \mathcal{S}_t). \quad (13)$$

We shall use h_t to represent a distorted expectation operator, conditioned on \mathcal{S}_t events, of random variables that are \mathcal{X}_t measurable. The random variable h_t can be used to build a joint density function for z_0, z_1, \dots, z_t conditioned on current and past signals. We can integrate this joint density to produce a marginal density for z_t conditioned on the signal history. This

⁶See Gelman, Carlin, Stern, and Rubin (1995), especially pages 276–283.

marginal density is relative to the density $q_t(z)$ that we constructed without distorting the underlying probability distribution. Thus, h_t implies a multiplicative perturbation of the hidden state density $q_t(z)$.

By exploiting the decomposition $M_t = h_t G_t$ and focusing directly on the distortion h_t , we can substantially streamline the problem of designing a robust \mathcal{S}_t -measurable decision rule for a_t . In particular, we can avoid having to solve a distorted filtering problem for each choice of M_t , a computationally demanding task that we would have to perform if we worked directly with M_t .⁷ By applying standard filtering and smoothing procedures to the approximating model, we can deduce $q_t(z)$, the probability of current and past hidden states given current signals implied by the benchmark model, then perturb it by multiplying by h_t . Because we do not have to resolve filtering problems for every potential perturbation, this can afford a considerable computational saving. When we have completed that task, we can then compute G_t and hence M_t .

The factorization $M_t = h_t G_t$ implies another interesting decomposition of discounted entropy. Hansen, Sargent, Turmuhambetova, and Williams (2004) condition on date zero information, which means that the time 0 contribution to entropy, $M_0 \log M_0$, plays no role in their analysis. However, the presence of hidden state variables will give $M_0 \log M_0$ an important role in our analysis. Including the $t = 0$ contribution and decomposing the martingale gives:

$$(1 - \rho) \sum_{t=0}^{\infty} \rho^t E(M_t \log M_t | \mathcal{S}_0) = (1 - \rho) \sum_{t=0}^{\infty} \rho^t E[G_t E(h_t \log h_t | \mathcal{S}_t) + G_t \log(G_t) | \mathcal{S}_0].$$

We will be primarily interested in the contribution:

$$(1 - \rho) \sum_{t=0}^{\infty} \rho^t E[G_t E(h_t \log h_t | \mathcal{S}_t) | \mathcal{S}_0] = (1 - \rho) \sum_{t=0}^{\infty} \rho^t E[G_t \varepsilon_t^2(h_t) | \mathcal{S}_0]$$

because of our ability to assess proposed current and future actions conditioned on the current signal history.

3.8 Two operators

We use the representations of entropy (8) and (13) to derive two operators that are useful for designing robust decision rules. Each of these operators can also be used to express a risk-sensitive adjustment of an appropriately measurable value function. Each operator also implies a distorted conditional expectation operator.

3.9 The operator R_t^1

Problem 3.10.

$$R_t^1(W_{t+1} | \theta) = \min_{m_{t+1} \in \mathcal{M}_{t+1}} E(m_{t+1} W_{t+1} | \mathcal{X}_t) + \theta \varepsilon_t^1(m_{t+1}).$$

⁷The \mathcal{S}_t -measurable random variable h_t distorts the distribution of \mathcal{X}_t . That implies that it distorts the conditional distribution $q_t(z)$ of z_t , i.e., the filter, and that it also distorts the distributions of z_0, \dots, z_{t-1} conditional on \mathcal{X}_t , i.e., the ‘smoothers’.

Claim 3.11. *Suppose that*

$$E \left[\exp \left(-\frac{W_{t+1}}{\theta} \right) \mid \mathcal{X}_t \right] < \infty. \quad (14)$$

The minimizing choice of m_{t+1} in problem 3.10 is

$$m_{t+1}^* \propto \exp \left(-\frac{W_{t+1}}{\theta} \right) \quad (15)$$

and the optimized value of the objective is

$$\mathbf{R}_t^1(W_{t+1}|\theta) = -\theta \log E \left[\exp \left(-\frac{W_{t+1}}{\theta} \right) \mid \mathcal{X}_t \right].$$

Proof. This result and variants of it that follow are standard in the literature on relative entropy. For example, see Dupuis and Ellis (1997). \square

In the limiting case that sets the entropy penalty parameter $\theta = \infty$, $\mathbf{R}_t^1(W_{t+1}|\infty) = E(W_{t+1}|\mathcal{X}_t)$. Notice that this expectation can depend on the hidden state. When $\theta < \infty$, \mathbf{R}_t^1 adjusts the ordinary continuation value $E(W_{t+1}|\mathcal{X}_t)$ 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 (15). When the conditional moment restriction (14) is not satisfied, we define \mathbf{R}_t^1 to be $-\infty$ on the relevant conditioning events.

3.10 The operator \mathbf{R}_t^2

For an \mathcal{X}_t -measurable function \hat{W}_t , the following problem implies an operator \mathbf{R}_t^2 and an associated worst-case distortion h :

Problem 3.12.

$$\mathbf{R}_t^2(\hat{W}_t|\theta) = \min_{h_t \in \mathcal{H}_t} E(h_t \hat{W}_t | \mathcal{S}_t) + \theta \varepsilon_t^2(h_t).$$

Claim 3.13. *Suppose that*

$$E \left[\exp \left(-\frac{\hat{W}_t}{\theta} \right) \mid \mathcal{S}_t \right] < \infty.$$

The minimizing choice of h_t in problem 3.12 is:

$$h_t^* \propto \exp \left(-\frac{\hat{W}_t}{\theta} \right)$$

$$\mathbf{R}_t^2(\hat{W}_t|\theta) = -\theta \log E \left[\exp \left(-\frac{\hat{W}_t}{\theta} \right) \mid \mathcal{S}_t \right].$$

4 Robust evaluations under commitment

This section states a robust decision problem under a timing protocol that requires each of two players in a zero sum game to choose appropriately measurable sequences of controls once and for all at time 0. We use outcomes of this timing protocol to motivate an alternative, sequential timing protocol in Hansen and Sargent (2005).

4.1 Full information

We want to study a problem in which a minimizing player has an information advantage (he conditions time t choices on \mathcal{X}_t) relative to a maximizing player (who conditions time t choices on \mathcal{S}_t). We obtain useful preliminary results by first considering a robust version of a full information decision problem. We represent a full information problem by temporarily replacing \mathcal{A}_t by an enlarged set \mathcal{A}_t^* that consists of all A -valued random vectors that are \mathcal{X}_t measurable.

This full information problem is the discrete-time counterpart to a continuous time problem of Hansen, Sargent, Turmuhambetova, and Williams (2004). The following two-player zero-sum game instructs a minimizing player to choose a martingale to perturb a maximizing player's probability model:

$$\max_{a_t \in \mathcal{A}_t^*} \min_{m_{t+1} \in \mathcal{M}_{t+1}} E \left(\sum_{t=0}^{\infty} M_t [\beta^t U(x_t, a_t) + \rho^t \theta m_{t+1} \log m_{t+1}] \mid \mathcal{X}_0 \right) \quad (16)$$

subject to

$$\begin{aligned} y_{t+1} &= \pi_y(s_{t+1}, y_t, a_t) \\ z_{t+1} &= \pi_z(x_t, a_t, w_{t+1}) \\ s_{t+1} &= \pi_s(x_t, a_t, w_{t+1}) \\ M_{t+1} &= m_{t+1} M_t \end{aligned} \quad (17)$$

where $M_0 = 1$ and x_0 is a known initial condition. We restrict both the distortion m_{t+1} and the decision a_{t+1} to be \mathcal{X}_{t+1} measurable.

The distortion in beliefs contributes a multiplicative martingale preference shock M_t to period- t utility and a θ -weighted penalty on entropy.⁸ The initial condition M_0 becomes a new state variable at time 0. In the full information case, we are free to normalize M_0 to one. A Bellman-Issacs condition allows us to exchange orders of maximization and minimization (see Hansen, Sargent, Turmuhambetova, and Williams (2004)).⁹

Remark 4.1. *The inner problem in game (16)–(17) is an ordinary control problem with a fully observed state (x_t, M_t) and an \mathcal{X}_t -measurable control m_{t+1} .*

While Hansen, Sargent, Turmuhambetova, and Williams (2004) study the case in which $\beta = \rho$, we also consider the case in which $0 < \beta < 1$ and $\rho = 1$. When $\rho = 1$ and the minimizing agent chooses first, we can solve the inner part of problem (16) as a static optimization problem. For a given action process $\{a_t : t \geq 0\}$, let

$$W_\infty \doteq \sum_{t=0}^{\infty} \beta^t U(x_t, a_t)$$

subject to

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

⁸Maccheroni, Marinacci, and Rustichini (2004) developed axiomatic treatments of penalty-based preferences. Wang (2001) developed axiomatic treatments of entropy-based measures.

⁹The freedom to exchange the order of extremization rationalizes a worst-case model under which the robust action for a_t is optimal (see Blackwell and Girshick (1954)).

$$\begin{aligned} z_{t+1} &= \pi_z(x_t, a_t, w_{t+1}) \\ s_{t+1} &= \pi_s(x_t, a_t, w_{t+1}). \end{aligned}$$

The minimizing agent solves:

Problem 4.2.

$$R_\infty^1(W_\infty) \doteq \min_{M_\infty \geq 0, E(M_\infty | \mathcal{X}_0) = 1} E(M_\infty W_\infty | \mathcal{X}_0) + \theta E(M_\infty \log M_\infty | \mathcal{X}_0).$$

Claim 4.3. For some action process $\{a_t : t \geq 0\}$, suppose that

$$E \left[\exp \left(-\frac{1}{\theta} W_\infty \right) | \mathcal{X}_0 \right] < \infty.$$

Then

$$M_\infty^* = \frac{\exp \left(-\frac{1}{\theta} W_\infty \right)}{E \left[\exp \left(-\frac{1}{\theta} W_\infty \right) | \mathcal{X}_0 \right]}$$

and the minimized value of the objective is:

$$R_\infty^1(W_\infty) = -\theta \log E \left[\exp \left(-\frac{1}{\theta} W_\infty \right) | \mathcal{X}_0 \right]. \quad (18)$$

The implied martingale is

$$M_t^* = E(M_\infty^* | \mathcal{X}_t),$$

and $M_0^* = 1$. Control theory interprets (18) as a risk-sensitive adjustment of the criterion W_∞ (e.g., see Whittle (1990)). The solution for the distortions $\{M_t^* : t \geq 0\}$ justifies an interpretation of that risk-sensitivity adjustment as expressing a concern about robustness. An expression of the outer problem in (16) is

$$\max_{a_t \in \mathcal{A}_t^*, t \geq 0} -\theta \log E \left(\exp \left[-\frac{1}{\theta} \sum_{t=0}^{\infty} \beta^t U(x_t, a_t) \right] \middle| \mathcal{X}_0 \right).$$

Setting $\beta < 1$ but $\rho = 1$ makes the concern about robustness wear off with the passage of time, in the sense that $\frac{M_{t+1}^*}{M_t^*} = m_{t+1}^* \rightarrow 1$ as $t \rightarrow +\infty$. We analyze a linear quadratic example in section 7.

Hansen and Sargent (1995), in discrete time, and Hansen, Sargent, Turmuhambetova, and Williams (2004), in continuous time, have formulated versions of this full-information problem when $\rho = \beta$. They analyzed a zero-sum two-player game in which both players choose sequentially for the case $\rho = \beta$. The Markov perfect equilibrium of the two-person zero-sum game gives rise to a value function for the maximizing player that can be stated in terms of a recursive version of an adjustment for risk-sensitivity. When $\rho = \beta$, the associated martingale is most conveniently represented in terms of continuation values. The martingale typically does not converge to a finite limiting random variable M_∞ when $\beta = \rho$. When it does not converge, concerns about robustness endure. When the martingale fails to converge, it can still imply a limiting probability distribution for \mathcal{X}_∞ , but that limiting distribution will typically is not absolutely continuous with respect to the distribution associated with the approximating model.

4.2 Hidden states

The problem in which the minimizing player has an information advantage entails two important modifications of the full information problem. First, although m_{t+1} remains \mathcal{X}_{t+1} measurable, the action a_t is restricted to be \mathcal{S}_t measurable and hence in \mathcal{A}_t rather than in the larger set \mathcal{A}_t^* . Second, M_0 now must be an \mathcal{X}_0 measurable random variable satisfying:

$$E(M_0|\mathcal{S}_0) = 1,$$

or $G_0 = 1$. This convenient normalization implies that we can replace the minimization over M_0 with a minimization over h_0 . Thus, the minimizing player's decision problem can be posed as:

$$\min_{h_0 \in \mathcal{H}_0} \min_{m_{t+1} \in \mathcal{M}_{t+1}, t \geq 0} E \left[\sum_{t=0}^{\infty} M_t \beta^t U(y_t, z_t, a_t) + \rho^t \theta m_{t+1} \log m_{t+1} | \mathcal{S}_0 \right] + \theta E(M_0 \log M_0 | \mathcal{S}_0) \quad (19)$$

subject to (17), a given action process whose time t component a_t is in \mathcal{A}_t , $M_0 = h_0$, and the observable component of the initial state y_0 as an initial condition. It is convenient to minimize over h_0 after first solving a preliminary inner problem¹⁰

$$W_0 \doteq \min_{m_{t+1} \in \mathcal{M}_{t+1}, t \geq 0} E \left(\sum_{t=0}^{\infty} M_t [\beta^t U(y_t, z_t, a_t) + \rho^t \theta m_{t+1} \log m_{t+1}] | \mathcal{X}_0 \right) \quad (20)$$

subject to (17) with $h_0 = 1$. Evidently, W_0 depends implicitly on the decision process $\{a_t : t \geq 0\}$ and on conditioning information in \mathcal{X}_0 .

Notice that problem (20)–(17) is identical with the full information problem for the minimizing agent (16)–(17). This inner problem has a recursive structure. In particular, the solution to problem (20) can be expressed in terms of successive applications of the operator \mathbf{R}_t^1 . Starting from a terminal value W_T that is \mathcal{X}_T measurable, compute W_t for $t = T - 1, T - 2, \dots, 0$ by performing the recursions

$$W_t = \frac{\beta^t}{\rho^t} U(y_t, z_t, a_t) + \rho \mathbf{R}_t^1(W_{t+1} | \theta), \quad (21)$$

where W_t is \mathcal{X}_t measurable. Think of solving these recursions as $T \rightarrow +\infty$. As a by-product of these recursions, compute the associated m_{t+1} process

$$m_{t+1} \propto \exp\left(\frac{-W_{t+1}}{\theta}\right), \quad t \geq 0. \quad (22)$$

The \mathbf{R}_t^1 operator applied to W_{t+1} yields a random vector m_{t+1} that distorts a one-step ahead density for x_{t+1} conditional on \mathcal{X}_t . The recursive solution strings together these conditional distortions to produce a worst-case distortion of \mathcal{X}_∞ events conditioned on the date zero information set \mathcal{X}_0 .

¹⁰In formulating problem (20)–(23), we are using the fact that the right side of (19) is linear in h_0 .

Remark 4.4. *Recursion (21) maps \mathcal{X}_{t+1} measurable random variables into \mathcal{X}_t measurable random variables. Because the spaces of histories \mathcal{X}_t are so large, these recursions are not computationally practical unless one can find a way to reduce the dimension of the state by, for example, deriving a sufficient statistic for histories. Section 6 describes the distorted distributions that emerge from these operators. In the linear quadratic examples of section 7, we display how to compute explicit examples of these distributions that are of reduced dimensions. Here an object $\check{q}_t(z)$ emerges that is a normal distribution with distorted mean and covariance that generalizes the ordinary Kalman filter. In a special case that one period utilities do not depend on the hidden state, $\check{q}_t(z)$ collapses to $q_t(z)$. Therefore, in that special case, the distribution $q_t(z)$ associated with the approximating model can serve as a state variable even though the minimizing agent distorts this distribution to attain a robust decision rule.*

For our incomplete information game, after solving the inner problem, we solve the *outer* problem:

$$\hat{W}_0 = \min_{h_0 \in \mathcal{H}_0} E(h_0 W_0 + \theta h_0 \log h_0 | \mathcal{S}_0). \quad (23)$$

The outer problem is identical with the problem that defines the operator R_0^2 . After having computed W_0 , we solve the outer problem by computing

$$\hat{W}_0 = R_0^2(W_0 | \theta).$$

The associated distortion h_0 can be computed from

$$h_0 \propto \exp\left(\frac{-W_0}{\theta}\right).$$

The application of the R_0^2 operator to W_0 completes the minimizing player's choice by yielding a random vector h_0^* that distorts the probabilities assigned to \mathcal{X}_0 conditioned on \mathcal{S}_0 . Thus, the minimization part of the commitment problem (19) is solved by first applying the R^1 operator 'an infinite number' of times, then applying the R^2 operator once, at time 0.

4.3 Important special cases

Substantial simplifications occur when we further restrict the problem. Many treatments of robustness in the control theory literature set $\rho = 1$. From problem (4.2), we know that when $\rho = 1$, there exists a random variable M_∞ that can be used to distort probabilities at all calendar dates. Problem (4.2) yields a choice of M_∞ that is \mathcal{X}_0 measurable. To solve for the worst-case model pertinent when there are hidden state variables, we would compute:

$$R_0^2[R_\infty^1(W_\infty)]$$

as in problem (19), where

$$W_\infty \doteq \sum_{t=0}^{\infty} \beta^t U(x_t, a_t).$$

Equivalently, the minimizing agent could solve:

Problem 4.5.

$$R_\infty^2(W_\infty) \doteq \min_{H_\infty \geq 0, E(H_\infty | \mathcal{S}_0) = 1} E(H_\infty W_\infty | \mathcal{S}_0) + \theta E(H_\infty \log H_\infty | \mathcal{S}_0).$$

The solution to this problem is:

$$H_\infty^* = \frac{W_\infty}{E(W_\infty | \mathcal{S}_0)} = \frac{E(W_\infty | \mathcal{X}_0)}{E(W_\infty | \mathcal{S}_0)} M_\infty^*.$$

where $M_\infty^* = \frac{W_\infty}{E(W_\infty | \mathcal{X}_0)}$ is the full information solution associated with R_∞^2 . We will illustrate such a solution in section 7 by appealing to results from the robust control theory literature that assume that $\rho = 1$.

When $\rho = 1$, we can decompose the random variable M_∞ as:

$$M_\infty = h_\infty G_\infty \tag{24}$$

where

$$G_\infty \doteq E(M_\infty | \mathcal{S}_\infty),$$

and

$$h_\infty \doteq \begin{cases} \frac{M_\infty}{G_\infty} & \text{if } G_\infty > 0 \\ 1 & \text{if } G_\infty = 0. \end{cases}$$

The random variable M_∞ distorts the probability distribution of \mathcal{X}_∞ , h_∞ distorts the probability of \mathcal{X}_∞ conditioned on \mathcal{S}_∞ , and G_∞ distorts the probability of \mathcal{S}_∞ . Use decomposition (24) to express entropy as:

$$E(M_\infty \log M_\infty | \mathcal{S}_0) = E[G_\infty E(h_\infty \log h_\infty | \mathcal{S}_\infty) | \mathcal{S}_0] + E[G_\infty \log G_\infty | \mathcal{S}_0].$$

We can put decomposition (24) to good use in the following special case:

Example 4.6. Suppose that $\rho = 1$ and that U does not depend on the hidden state z_t :

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

Because W_∞ is \mathcal{S}_∞ measurable, the minimizing solution to problem (4.5) sets $h_\infty = 1$ and

$$G_\infty = \frac{W_\infty}{E(W_\infty | \mathcal{S}_0)}.$$

Although the minimizing agent has an informational advantage in this example, he does not use it in the limit (he sets $h_\infty = 1$). The worst case $M_\infty = G_\infty$ distorts only the signal distribution and not the distribution of the states conditioned on the entire signal process. However, it will distort the distribution of z_t conditioned on \mathcal{S}_t because, while M_∞ will be \mathcal{S}_∞ measurable, $M_t = E(M_\infty | \mathcal{X}_t)$ will typically not be \mathcal{S}_t measurable. There will be events in \mathcal{X}_t with probability assignments conditioned on \mathcal{S}_t that are distorted. In particular, the distortion of the conditional expectation of a real valued, bounded Borel measurable function ϕ of z_t given \mathcal{S}_t , $\frac{E[M_t \phi(z_t) | \mathcal{S}_t]}{E[M_t | \mathcal{S}_t]}$, can be different from $E[\phi(z_t) | \mathcal{S}_t]$.

Of course, distortions of probabilities over events that are measurable with respect to the filtration $\{\mathcal{S}_t : t \geq 0\}$, not the implied distortions of the distribution of z_t conditioned on \mathcal{S}_t , are the pertinent distortions for the choice of action process $\{a_t : t \geq 0\}$.

4.4 An alternative approach for Example 4.6

When $\rho = 1$, we can use the following decomposition of the entropy of the distortion G_∞ :

$$E [G_\infty \log G_\infty | \mathcal{S}_0] = \sum_{j=0}^{\infty} E [G_j E (g_{j+1} \log g_{j+1} | \mathcal{S}_j) | \mathcal{S}_0] \quad (25)$$

where g_{j+1} for $j \geq 0$ is constructed from:

$$g_{j+1} = \begin{cases} \frac{G_{j+1}}{G_j} & \text{if } G_j > 0 \\ 1 & \text{if } G_j = 0 \end{cases}$$

An equivalent statement of the minimizing agent's problem is

$$\min_{g_{t+1} \in \mathcal{G}_{t+1}, t \geq 0} E \left(\sum_{t=0}^{\infty} G_t \left[\beta^t \hat{U}(y_t, a_t) + \theta g_{t+1} \log g_{t+1} \right] | \mathcal{S}_0 \right) \quad (26)$$

subject to

$$\begin{aligned} y_{t+1} &= \pi_y(s_{t+1}, y_t, a_t) \\ z_{t+1} &= \pi_z(x_t, a_t, w_{t+1}) \\ s_{t+1} &= \pi_s(x_t, a_t, w_{t+1}) \\ G_{t+1} &= g_{t+1} G_t \end{aligned} \quad (27)$$

where $G_0 = 1$ and \mathcal{G}_{t+1} consists of the set of all nonnegative \mathcal{S}_{t+1} measurable random variables with expectation conditioned on \mathcal{S}_t equal to unity. This takes the form of an ordinary (i.e., nonrobust) control problem. In this formulation, there is no informational advantage for the minimizing agent and therefore no initial minimization in setting the initial conditions for the martingale $\{G : t \geq 0\}$ adapted to the smaller filtration $\{\mathcal{S}_t : t \geq 0\}$. A counterpart to the R_t^1 operator can be constructed by replacing \mathcal{X}_t with \mathcal{S}_t and m_{t+1} with g_{t+1} . The date t operator will necessarily use θ/β^t as the date t robustness parameter.

The information structure and the structure of the constraints in problem (26)-(27) invite the minimizing player to solve a filtering problem, just as in the standard control problem with hidden states. Thus, by the same logic that connects problems 2.1 and 2.2, problem (26)-(27) could be restated as

$$\min_{g_{t+1} \in \mathcal{G}_{t+1}, t \geq 0} E \left(\sum_{t=0}^{\infty} G_t \left[\beta^t \hat{U}(y_t, a_t) + \theta g_{t+1} \log g_{t+1} \right] | \mathcal{S}_0 \right)$$

subject to the density $\varsigma(s_{t+1} | y_t, q_t, a_t)$ and

$$\begin{aligned} y_{t+1} &= \pi_y(s_{t+1}, y_t, a_t) \\ q_{t+1} &= \pi_q(s_{t+1}, y_t, q_t, a_t) \\ G_{t+1} &= g_{t+1} G_t. \end{aligned}$$

The g_{t+1} choice will distort the transition probabilities for \mathcal{S}_{t+1} events given \mathcal{S}_t information. The implied distortions for \mathcal{X}_{t+1} conditioned on \mathcal{X}_t could be computed from the

formula:

$$m_{t+1} = \frac{E \left(\exp \left[-\frac{\beta^t}{\theta} \sum_{j=0}^{\infty} \beta^j \hat{U}(y_{t+j}, a_{1+j}) \right] \mid \mathcal{X}_{t+1} \right)}{E \left(\exp \left[-\frac{\beta^t}{\theta} \sum_{j=0}^{\infty} \beta^j \hat{U}(y_{t+j}, a_{t+j}) \right] \mid \mathcal{X}_t \right)},$$

but it is not necessary to perform this calculation to determine the maximizing actions. Notice that when $\rho = 1$, the implicit date t robustness parameter is $\frac{\theta}{\beta^t}$. Therefore, when $0 < \beta < 1$, the impact of the concern about robustness vanishes with the passage of time, a point that was noted, for instance, by Whittle (1981).

5 Relationship to Chamberlain's and Knox's work

Before studying example 4.6 in the next section, it is useful to digress briefly to consider how the robust estimation problems studied by Chamberlain (2000) and Knox (2003) fit into our framework, especially in terms of the behavior of the martingale distortion M_t .

Example 5.1. *Consider an estimation problem in which $z_t = z_0$ for all t . The hidden state does not evolve over time, and as consequence, can be viewed as an unknown parameter to be estimated. Signals provide information about this parameter.*

Suppose that \mathcal{X}_t is generated by \mathcal{S}_t and z_0 . Thus M_t depends on the signal history and z_0 , and, similarly, m_{t+1} depends on the date $t+1$ signal, the signal history, and z_0 . The random variable m_{t+1} distorts the conditional distribution of s_{t+1} given current and past signals and the unknown parameter. The random variable h_t distorts the posterior distribution of the parameter $z_0 = z_t$ conditioned on \mathcal{S}_t . While signal distributions can be distorted, absolute continuity requires the distorted distribution of z_t to be invariant for all $t \geq 0$.

The ability to distort the signal distribution conditioned on the parameter differentiates this problem from those studied by Chamberlain and Knox. They study a robust parameter estimation problem, impose commitment, and make a robust choice of a *prior* M_0 , where $E(M_0 | \mathcal{S}_0) = 1$. In effect, they restrict the martingale $\{M_t : t \geq 0\}$ to be time invariant. The implied form of their h_t distortion is thus:

$$h_t = \frac{M_0}{E(M_0 | \mathcal{S}_t)}$$

rather than our less restrictive specification.

We believe that prior selection problems like Chamberlain (2000) and Knox (2003) are of considerable interest. But those problems differ substantially from ours because we allow distortions of the signal distribution, given the parameter, and do not limit the martingale $\{M_t : t \geq 0\}$ to be time invariant.

6 Recursive formulation of commitment game with undiscounted entropy

When $\rho = 1$, we can solve the commitment problem by simply finding a worst-case M_∞ that distorts probabilities assigned to events in \mathcal{F}_∞ :

$$M_\infty = \frac{\exp \left[-\frac{1}{\theta} \sum_{t=0}^{\infty} \beta^t U(x_t^*, a_t^*) \right]}{E \left(\exp \left[-\frac{1}{\theta} \sum_{t=0}^{\infty} \beta^t U(x_t^*, a_t^*) \right] \mid \mathcal{S}_0 \right)}.$$

The $*$ variables denote objects evaluated at the commitment solution.

We now describe a recursive formulation of the $\rho = 1$ commitment problem by characterizing a corresponding date k preference ranking.¹¹ The idea will be to find a way to ‘decentralize the decisions of the minimizing agent over time’ by having a sequence of minimizing players who take as given the decisions of past minimizing agents and the decision rules of future minimizing agents. In addition to leading to simplified computational algorithms, the depiction of a conditional ranking allows us to isolate the distinct roles played by commitment and discounting. We view this analysis as a stepping stone to the formulations of recursive games in Hansen and Sargent (2005) that withdraw the ability to commit from the two players in a zero-sum game designed to produce a robust decision rule.

6.1 Date k probability distortion

As a useful prolegomenon to constructing a date k conditional preference ordering, we introduce the following decomposition of M_∞ into backward- and forward-looking components:

$$M_\infty = M_{k-1}^b M_\infty^f$$

where

$$\begin{aligned} M_{k-1}^b &\doteq \frac{\exp \left[-\frac{1}{\theta} \sum_{t=0}^{k-1} \beta^t U(x_t^*, a_t^*) \right]}{E \left(\exp \left[-\frac{1}{\theta} \sum_{t=0}^{k-1} \beta^t U(x_t^*, a_t^*) \right] \mid \mathcal{S}_0 \right)} \\ M_\infty^f &\doteq \exp \left[-\frac{1}{\theta} \sum_{t=k}^{\infty} \beta^t U(x_t^*, a_t^*) \right] \frac{E \left(\exp \left[-\frac{1}{\theta} \sum_{t=0}^{k-1} \beta^t U(x_t^*, a_t^*) \right] \mid \mathcal{S}_0 \right)}{E \left(\exp \left[-\frac{1}{\theta} \sum_{t=0}^{\infty} \beta^t U(x_t^*, a_t^*) \right] \mid \mathcal{S}_0 \right)}. \end{aligned}$$

The term M_{k-1}^b is backward looking and therefore depends on utility contributions before date k . Changes in actions from date k forward will not alter this component. The corresponding entropy decomposition is:

$$E(M_\infty \log M_\infty \mid \mathcal{S}_0) = E(M_\infty^f M_{k-1}^b \log M_{k-1}^b \mid \mathcal{S}_0) + E(M_\infty^f M_{k-1}^b \log M_\infty^f \mid \mathcal{S}_0)$$

Construct two martingales

$$M_t = E(M_\infty \mid \mathcal{X}_t) = M_{k-1}^b E(M_\infty^f \mid \mathcal{X}_t)$$

¹¹Although the mathematical formulation differs in important ways because we use martingales to represent distorted models, our arguments parallel some of those made in the robust control literature by Whittle (1990) and Basar and Bernhard (1995).

and

$$M_t^f = E(M_\infty^f | \mathcal{X}_t).$$

The two martingales share increments for $t \geq k$:

$$m_{t+1} = \frac{M_{t+1}}{M_t} = \frac{M_{t+1}^f}{M_t^f}.$$

6.2 Changing the benchmark probability distribution at date k

We use M_{k-1}^b to define a new benchmark model at date k , then explore distortions to that model by choosing a worst case M_∞^f . The notation \check{E} denotes an expectation computed with a distorted probability measure $\check{P}r$ constructed by taking M_{k-1}^b to be a Radon-Nikodym derivative with respect to the probability distribution associated with the original (time 0) approximating model. Write the date zero objective evaluated at the commitment solution as:

$$E \left[M_\infty \sum_{t=0}^{\infty} \beta^j U(x_t^*, a_t^*) | \mathcal{S}_0 \right] = \check{E} \left[M_\infty^f \sum_{t=0}^{k-1} \beta^j U(x_t^*, a_t^*) | \mathcal{S}_0 \right] + \check{E} \left[M_k^f \sum_{t=k}^{\infty} \beta^t M_t^k U(x_t^*, a_t^*) | \mathcal{S}_0 \right].$$

Notice that

$$\begin{aligned} \theta E (M_\infty^f M_{k-1}^b \log M_{k-1}^b | \mathcal{S}_0) &= - \check{E} \left[M_k^f \sum_{t=0}^{k-1} \beta^j U(x_t^*, a_t^*) | \mathcal{S}_0 \right] \\ &\quad + \log \left[E \left(\exp \left[-\frac{1}{\theta} \sum_{t=0}^{k-1} \beta^t U(x_t^*, a_t^*) \right] | \mathcal{S}_0 \right) \right]. \end{aligned}$$

Thus, including the entropy penalty, the date zero objective can be rewritten as:

$$\begin{aligned} \check{E} \left[M_k^f \sum_{t=k}^{\infty} \beta^t M_t^k U(x_t^*, a_t^*) | \mathcal{S}_0 \right] &+ \theta \check{E} [M_\infty^f \log M_\infty^f | \mathcal{S}_0] \\ &+ \log \left[E \left(\exp \left[-\frac{1}{\theta} \sum_{t=0}^{k-1} \beta^t U(x_t^*, a_t^*) \right] | \mathcal{S}_0 \right) \right]. \quad (28) \end{aligned}$$

We will exploit this representation to implement the commitment problem recursively. Formally, fixing the distorted expectation operator \check{E} at date k , we explore changes in the control policy and penalized choices of worst-case models relative to the $\check{P}r$ probability distribution. This allows us to work with entropy relative to the $\check{P}r$ probability distribution and to let the objective include terms only from date k forward. From the solution to Problem 3.8, we know that we can solve a conditional counterpart to this problem in which we decompose:

$$M_k^f = h_k^f G_k^f$$

with G_k^f being \mathcal{S}_k measurable and h_k^f satisfying:

$$\check{E}(h_k^f | \mathcal{S}_k) = 1.$$

The conditional problem uses the objective:

$$\check{E} \left[M_k^f \sum_{t=k}^{\infty} \beta^t M_t^k U(x_t, a_t) | \mathcal{S}_k \right] + \theta \check{E} \left[h_k^f \log(h_k^f) | \mathcal{S}_k \right] + \check{E} \left(h_k^f M_\infty^k \log(M_\infty^k) | \mathcal{S}_k \right).$$

and abstracts from the minimizing choice of G_k^f . The third term in (28) is irrelevant for this conditional problem.

We shall use the notation $\check{\mathcal{H}}_k$ to denote the analog to \mathcal{H}_k constructed using the $\check{P}r$ probability measure in place of the original probability measure.

6.3 Reconsidering the date k decision

To induce a recursive algorithm for solving the commitment problem, we divide the maximizing player into a sequence of date k maximizing players and ask each date k player to reconsider the date k choice that our single maximizing player chose under commitment, taking as given the contingency rules of the maximizing agents for dates $s \neq k$. If we endow each date k maximizing agent with the same date zero objective that we used under commitment, the date k player will reaffirm the date k component chosen under commitment. However, to get a recursive implementation of the commitment solution, we want the date k maximizing decision-maker to use an objective based only on one-period utilities from date k forward, i.e., we want him to consult continuation values. In addition to ignoring earlier contributions to the objective, we would also ask an associated time t *minimizing* agent to choose h_t .

In this section, we shall eventually describe the rankings over action processes implied by minimization in terms of applications of our two operators R_1 and R_2 . We construct date k minimization problems that preserve conditional rankings for any $k \geq 0$ and thereby motivate dynamic programming formulations. We shall represent x_t as a function of current and past actions and shocks

$$x_t = \psi_t(a_0, a_1, \dots, a_{t-1}),$$

leaving implicit the dependence of the random function ψ_t on the underlying shocks and the initial state.

We now rank contingent decision processes from date k forward. We denote the infinite sequence of current and future decisions by

$$a_{k,\infty} \doteq (a_k, a_{k+1}, a_{k+2}, \dots)$$

and define

$$\xi_t(a_{k,\infty}) \doteq \psi_t(a_0^*, a_1^*, \dots, a_{k-1}^*, a_k, \dots, a_{t-1})$$

for $t > k$. For $\ell \geq k$, let

$$\Xi_\ell(a_{k,\infty}) \doteq \sum_{t=\ell}^{\infty} \beta^t U[\xi_t(a_{k,\infty}), a_t]. \quad (29)$$

To produce a ranking over action process $a_{k,\infty}$, solve:

$$\min_{h_k \in \check{\mathcal{H}}_k} \min_{m_t \in \mathcal{M}_t, t > k} \check{E} \left[h_k M_\infty^k \Xi_k(a_{k,\infty}) + \theta h_k M_\infty^k \log M_\infty^k + \theta h_k \log h_k | \mathcal{S}_k \right]$$

where

$$M_\infty^k = \prod_{t>k}^{\infty} m_t.$$

Consider first the inner minimization problem:

$$\Xi_k^*(a_{k,\infty}) = \min_{m_t \in \mathcal{M}_t, t>k} E [M_\infty^k \Xi_k(a_{k,\infty}) | \mathcal{X}_k] + \theta E (M_\infty^k \log M_\infty^k | \mathcal{X}_k). \quad (30)$$

In section 4.2 in equation (21), we saw that the minimization over M_∞^k can be done recursively by iterating on

$$\Xi_\ell^*(a_{k,\infty}) = \beta^\ell U [\xi_\ell(a_{k,\infty}), a_\ell] + \mathbf{R}_\ell^1 [\Xi_{\ell+1}^*(a_{k,\infty}) | \theta], \quad (31)$$

coming backwards to $\ell = k$.

We can continue to use the \mathbf{R}_ℓ^1 operator because the $\check{P}r$ probability measure does not distort probabilities conditioned on \mathcal{X}_k . We use

$$\check{\mathbf{R}}_k^2 [\Xi_k^*(a_{k,\infty}) | \theta]$$

to rank alternative action processes, where $\check{\mathbf{R}}_k^2$ is constructed just as \mathbf{R}_k^2 but with the $\check{P}r$ probability measure in place of the original probability measure.

Claim 6.1. *Suppose that $\{a_t^* : t \geq 0\}$ solves the commitment problem. Then*

$$\check{\mathbf{R}}_k^2 [\Xi_k^*(a_{k,\infty}^*) | \theta] \geq \check{\mathbf{R}}_k^2 [\Xi_k^*(a_{k,\infty}) | \theta] \quad (32)$$

for any other feasible action process $a_{k,\infty}$, where $\Xi_k^*(a_{k,\infty})$ solves (30).

To represent the solution of the commitment problem recursively, we must somehow make the time k minimizer behave as the time 0 minimizer does under commitment. We accomplish this by constructing the $\check{P}r$ measure. The effects on the time k minimizer's problem of the terms in the objective preceding date k are confined to the $\check{P}r$ probability measure. To make the time k agent act like a time 0 agent under commitment, we require the time k agent to act as if the distorted measure $\check{P}r$ is his approximating model. Doing this imposes on the time k minimizing agent a commitment to the distorted measure $\check{P}r$ that imparts a form of consequentialism to the decision problem. However, we will allow the time k decision maker to contemplate additional probability distortions captured by M_∞^k, h_k . We will say more about the distortion embedded in the time k decision maker's 'benchmark model' and its consequences in later subsections.

6.4 The rear view mirror: distorted and undistorted filtering

The distortion in the benchmark model requires that we compute conditional expectations of the form

$$\check{E} [\phi(z_k) | \mathcal{S}_k].$$

for Borel measurable functions ϕ . Equivalently, we require the $\check{P}r$ probability conditional density \check{q}_k of z_k (relative to λ) conditioned on \mathcal{S}_k . When the utility function does not depend on the hidden state as in Example 4.6, M_k^b is \mathcal{S}_k measurable and the \check{q}_k probability distribution agrees with the conditional distribution under the original probability model.

But more generally, the distribution \check{q}_k can be computed recursively by adapting standard methods that entail *forward induction*. Begin by forming the multiplicative decomposition:

$$M_{k-1}^b = \frac{\prod_{t=0}^{k-1} \check{m}_t}{E\left(\prod_{t=0}^{k-1} \check{m}_t \mid \mathcal{S}_0\right)}$$

where

$$\check{m}_t \doteq \exp\left[-\frac{1}{\theta}\beta^t U(x_t^*, a_t^*)\right]. \quad (33)$$

Notice that M_k^b does not distort probabilities conditioned on \mathcal{X}_{k-1} . Therefore, given a density \check{q}_{k-1} (relative to the measure λ) of z_{k-1} conditional on \mathcal{S}_{k-1} , we can infer \check{q}_{k-1} using:

$$\check{q}_k = \pi_q(s_k^*, y_{k-1}^*, \check{q}_{k-1}, a_{k-1}^*).$$

It remains to compute \check{q}_{k-1} .

This sequence of densities $\{\check{q}_k\}$ can be computed recursively by constructing a sequence of probability models. Form a family of distorted probability measures $\{\widetilde{P}r_j : j \geq 0\}$ in which measure $\widetilde{P}r_j$ has Radon-Nikodym derivative $\widetilde{M}_j = M_j^b$. We thereby construct these measures so that $\widetilde{P}r_{k-1}$ coincides with $\check{P}r$. Let \tilde{q}_j be the density (relative to the measure λ) induced by $\widetilde{P}r_j$ for z_j . Then

$$\check{q}_{k-1} = \tilde{q}_{k-1}.$$

To produce an updating equation for the density sequence $\{\tilde{q}_j : j \geq 0\}$, express the dependence of \check{m}_j on the hidden state by

$$\check{m}_j = \nu_j(z_j),$$

and note that

$$\begin{aligned} \tilde{q}_j(z) &= \frac{\nu_j(z)\pi_q(s_j^*, y_{j-1}^*, \tilde{q}_{j-1}, a_{j-1}^*)(z)}{\int \nu_j(\tilde{z})\pi_q(s_j^*, y_{j-1}^*, \tilde{q}_{j-1}, a_{j-1}^*)(\tilde{z})d\lambda(\tilde{z})} \\ &= \frac{\nu_j(z) \int \tau(z, s_j^* | y_{j-1}^*, \hat{z}, a_{j-1}^*) \tilde{q}_{j-1}(\hat{z}) d\lambda(\hat{z})}{\int \int \nu_j(\tilde{z}) \tau(\tilde{z}, s_j^* | y_{j-1}^*, \hat{z}, a_{j-1}^*) \tilde{q}_{j-1}(\hat{z}) d\lambda(\hat{z}) d\lambda(\tilde{z})}. \end{aligned}$$

This operation *updates* using π_q and *distorts* using the conditional density ν_j scaled appropriately. The recursion is initialized by scaling the initial density q_0 by $\nu_0(z)$, then renormalizing so that the product integrates to one.

6.5 Discounting

It is revealing to adjust for discounting in forming continuation values. We accomplish this by first multiplying Ξ_ℓ defined in (29) for $\ell \geq k$ by $\beta^{-\ell}$ to obtain:

$$\tilde{\Xi}_\ell(a_{k,\infty}) \doteq \beta^{-\ell} \Xi_\ell(a_{k,\infty}) = \sum_{t=0}^{\infty} \beta^t U[\xi_{t+\ell}(a_{k,\infty}), a_{t+\ell}].$$

Multiplying both sides of (30) by β^{-k} results in:

$$\tilde{\Xi}_k^*(a_{k,\infty}) = \beta^{-k} \Xi_k^*(a_{k,\infty}) = \min_{m_t \in \mathcal{M}_t, t > k} E \left[M_\infty^\ell \tilde{\Xi}_k(a_{k,\infty}) | \mathcal{X}_\ell \right] + \beta^{-k} \theta E \left(M_\infty^\ell \log M_\infty^\ell | \mathcal{X}_\ell \right).$$

It follows that we can calculate $\tilde{\Xi}_k^*$ recursively by iterating on

$$\tilde{\Xi}_\ell^*(a_{\ell,\infty}) = U[\xi_\ell(a_{k,\infty}), a_\ell] + \beta R_\ell^1(\tilde{\Xi}_{\ell+1}^*(a_{\ell+1,\infty}) | \beta^{-\ell-1} \theta) \quad (34)$$

backwards on ℓ to $k+1$. Similarly, we use

$$\check{R}^2[\tilde{\Xi}_k^* | \beta^{-k} \theta] \quad (35)$$

to rank the alternative action processes.

Thus, the standard approach of renormalizing the discount factor at date k to be unity can be accomplished only by introducing time dependence into the robustness parameter. Scaling θ by β^{-k} increases the robustness penalty geometrically in k and causes concern about robustness to wear off over time when β is less than one. This wearing off occurs because we discount the return function but not the entropy measure.

In Hansen and Sargent (2005), we show how to arrest that wearing off by using with θ to replace $\beta^{-k} \theta$ in recursions closely related to (34) and (35).

6.6 Example 4.6 revisited

When the special circumstances of example 4.6 prevail, there is available a second approach to producing a conditional ranking that avoids solving a filtering problem by working directly with the signal process and using the fact $h_\infty = 1$ for feasible action processes. Thus, assume that hidden states do not occur as arguments of U so that it can be written $U(y_k, a_k)$ at date k . Construct

$$G_\infty^k = \prod_{t=k+1}^{\infty} g_t.$$

Let the minimizing agent solve:

$$\Xi_k^*(a_{k,\infty}) = \min_{g_t \in \mathcal{G}_t, t \geq k+1} E \left[G_\infty^k \Xi_k(a_{k,\infty}) | \mathcal{X}_k \right] + \theta E \left(G_\infty^k \log G_\infty^k | \mathcal{S}_k \right)$$

Construct a conditional ranking of action processes by using

$$\beta^k \hat{U}(y_k^*, a_k) + \Xi_k^*(a_{k,\infty}). \quad (36)$$

Claim 6.2. *Suppose that $\{a_t^* : t \geq 0\}$ solves the commitment problem. Then*

$$\beta^k \hat{U}(y_k^*, a_k^*) + \Xi_k^*(a_{k,\infty}^*) \geq \beta^k \hat{U}(y_k^*, a_k) + \Xi_k^*(a_{k,\infty})$$

for any other feasible $a_{k,\infty}$ where $\Xi_k^*(a_{k,\infty})$ solves (36).

This approach works with the signal filtration from the outset and does not directly assign distorted probabilities to hidden states (i.e., it sidesteps filtering problems). After the problem is solved, we can construct $M_\infty = G_\infty$ and deduce the implied probabilities for \mathcal{X}_k for any choice of k .

7 A linear quadratic decision problem under commitment

This section studies a decision problem with linear transition laws and a quadratic return function. We first solve a version of the problem under commitment, thereby implementing a particular version of the structure in section 4. Under commitment, Basar and Bernhard (1995) and Whittle (1990) have studied both undiscounted and discounted versions of this linear-quadratic problem and have displayed an elegant solution that incorporates a *worst-case certainty equivalent principle*. The certainty equivalence property permits one to solve the problem by combining solutions to two separate problems, namely, a nonstochastic control problem and an estimation problem. We display versions of these two subproblems in the following subsections. Then we relate our martingale representation of perturbations to formulas from the control theory literature.

The linear quadratic structure provides a practical context for us to expand on the fact that the equilibrium of a two-player zero-sum game under commitment has a recursive representation and therefore can be computed recursively. In addition, this setting enables a simple characterization of the probability distribution $\check{P}r$ that summarizes the history of signals and that functions as part of the time k approximating model in a recursive representation of the equilibrium of the game under commitment. In particular, we can summarize $\check{P}r$ by a particular Gaussian distribution $\check{q}_t(z)$ with mean \check{z} and covariance $\check{\Sigma}_k$. For the commitment problem, we describe the important special case in which $\check{q}_t(z) = q_t(z)$, where $q_t(z)$ is now the distribution associated with the ordinary Kalman filter.

7.1 The problem

The limited information problem under commitment is

$$\begin{aligned}
 - \max_{\{a_t \in \mathcal{A}_t\}} \min_{h_0 \in \mathcal{H}_0} \min_{\{m_{t+1} \in \mathcal{M}_{t+1}\}} \frac{1}{2} & E \sum_{t=0}^{\infty} M_t \left(\beta^t \begin{bmatrix} a_t' & x_t' \end{bmatrix} \begin{bmatrix} Q & P \\ P' & R \end{bmatrix} \begin{bmatrix} a_t \\ x_t \end{bmatrix} + \theta m_{t+1} \log m_{t+1} | \mathcal{S}_0 \right) \\
 & + \theta E(h_0 \log h_0 | \mathcal{S}_0) \tag{37}
 \end{aligned}$$

subject to

$$\begin{aligned}
 x_{t+1} &= Ax_t + Ba_t + Cw_{t+1} \\
 s_{t+1} &= Dx_t + Gw_{t+1} \\
 M_{t+1} &= m_{t+1}M_t \\
 M_0 &= h_0 \tag{38}
 \end{aligned}$$

where a_t is restricted to be measurable with respect to the sigma algebra \mathcal{S}_t . We assume that GG' is nonsingular and is independent of \mathcal{X}_t and that w_{t+1} has a multivariate, standard normal distribution. The decision maker fears that the conditional Gaussian distributions in (38) are misspecified.

To relate this to our earlier setup, we partition x_t as

$$x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}.$$

In section 2, we specified an evolution equation π_y for y_{t+1} and a joint density τ for the hidden state z_{t+1} and the signal s_{t+1} conditioned on x_t . The specification for τ can be obtained directly from (38) as multivariate normal. Suppose that π_y is chosen to satisfy

$$y_{t+1} = \Pi_s s_{t+1} + \Pi_y y_t + \Pi_a a_t.$$

Substituting from the evolution equation for the signal in (38), we obtain:

$$y_{t+1} = \Pi_s D x_t + \Pi_y y_t + \Pi_a a_t + \Pi_s G w_{t+1},$$

which gives the y -rows in system (38).

7.2 Full information without random shocks

The first step is to solve the following two-player zero-sum game with no uncertainty. The problem is

$$\max_{\{a_t\}} \min_{\{v_t\}} -\frac{1}{2} \sum_{t=0}^{\infty} \left(\beta^t \begin{bmatrix} a_t' & x_t' \end{bmatrix} \begin{bmatrix} Q & P \\ P' & R \end{bmatrix} \begin{bmatrix} a_t \\ x_t \end{bmatrix} + \theta |v_t|^2 \right) \quad (39)$$

subject to

$$x_{t+1} = A x_t + B a_t + C v_t \quad (40)$$

This deterministic problem follows an important part of the robust control theory literature in interpreting the ‘shocks’ v_t as unknown model misspecifications. As much of the robust control theory literature also does, problem (39)-(40) eliminates randomness. In section 7.3, we shall interpret v_t as a component of entropy in a random counterpart to this problem. Notice that we in effect set $\rho = 1$ in (16) and so do not discount the entropy term v_t . An action a_t can be made contingent on the state vector x_t .

We are interested in the equilibrium of the zero sum game (39)-(40) in which both players choose once and for all at time 0, i.e., a game in which both players commit to sequences of possibly state-contingent decisions at time 0. In more general settings, timing protocols affect equilibrium outcomes. However, in this particular zero-sum, full information game, it is known that identical equilibrium outcomes prevail across a set of games with alternative timing protocols: the details of the timing protocols affect neither equilibrium outcomes nor the recursive representations of equilibrium strategies. Therefore, in the present context, we lose no insights by focusing on a Markov perfect equilibrium in which both the maximizing and the minimizing decision makers choose sequentially.¹²

For sufficiently large values of θ , the Markov perfect equilibrium gives rise to a date t value function that is quadratic. Inclusive of discounting,¹³ we denote it

$$-\frac{\beta^t}{2} (x_t)' \Omega_t x_t.$$

¹²See Hansen, Sargent, Turmuhambetova, and Williams (2004) and Hansen and Sargent (2004a) for statements and proofs of the equivalence of outcomes and representations for different timing protocols.

¹³This problem is well posed only for sufficiently large values of θ . See Lemma 3.1 of Basar and Bernhard (1995).

For notational convenience, define:

$$\begin{aligned}\tilde{Q}_t &\doteq \begin{bmatrix} Q & 0 \\ 0 & -\beta^{-t}\theta I \end{bmatrix} \\ \tilde{P} &\doteq \begin{bmatrix} P \\ 0 \end{bmatrix} \\ \tilde{R} &\doteq R - \tilde{P}'(\tilde{Q}_t)^{-1}\tilde{P} = R - P'Q^{-1}P \\ \tilde{B} &\doteq \begin{bmatrix} B & C \end{bmatrix} \\ \tilde{A} &\doteq A - \tilde{B}(\tilde{Q}_t)^{-1}\tilde{P} = A - BQ^{-1}P.\end{aligned}$$

The robust a_t and the worst-case v_t are given by:

$$\begin{aligned}\begin{bmatrix} a_t \\ v_t \end{bmatrix} &= - \begin{bmatrix} Q + B'\beta\Omega_{t+1}B & \beta B'\Omega_{t+1}C \\ \beta C'\Omega_{t+1}B & \beta C'\Omega_{t+1}C - \beta^{-t}\theta I \end{bmatrix}^{-1} \begin{bmatrix} \beta B'\Omega_{t+1}A + P \\ \beta C'\Omega_{t+1}A \end{bmatrix} x_t \\ &= - \left(\beta [\tilde{Q}_t + \beta \tilde{B}'\Omega_{t+1}\tilde{B}]^{-1} \tilde{B}'\Omega_{t+1}\tilde{A} + (\tilde{Q}_t)^{-1}\tilde{P} \right) x_t,\end{aligned}\quad (41)$$

where the matrix Ω_t in the value function satisfies the Riccati equation

$$\Omega_t = \tilde{R} + \beta \tilde{A}'\Omega_{t+1}\tilde{A} - \beta \tilde{A}'\Omega_{t+1}\tilde{B} [\tilde{Q}_t + \beta \tilde{B}'\Omega_{t+1}\tilde{B}]^{-1} \tilde{B}'\Omega_{t+1}\tilde{A}.\quad (42)$$

(See Basar and Bernhard (1995) page 272.)

When $\beta < 1$, as $t \rightarrow +\infty$, the solution for Ω_t converges to the one that would be obtained under a no-robustness ($\theta = \infty$) specification, v_t converges to zero, and the limiting control law converges to that associated with $\theta = \infty$. When $\theta < +\infty$, the decision maker is concerned about robustness, but that concern diminishes over time. The dissipation of concerns about robustness with the passage of time is a direct consequence of the different discounting of one-period returns (they are discounted) and one-period entropies (they are not discounted).

7.3 Full information with random shocks

When random shocks are included, we are interested in solving the two-player zero-sum game¹⁴

$$-\frac{1}{2} \max_{\{a_t \in \mathcal{A}_t^*\}} \min_{M_\infty > 0, E(M_\infty | \mathcal{X}_\infty) = 1} E \left[M_\infty \sum_{t=0}^{\infty} \left(\beta^t \begin{bmatrix} a_t' & x_t' \end{bmatrix} \begin{bmatrix} Q & P \\ P' & R \end{bmatrix} \begin{bmatrix} a_t \\ x_t \end{bmatrix} \right) | \mathcal{X}_0 \right] + \theta E(M_\infty \log M_\infty | \mathcal{X}_0)\quad (43)$$

subject to:

$$x_{t+1} = Ax_t + Ba_t + Cw_{t+1}.\quad (44)$$

The date t admissible control set \mathcal{A}_t^* consists of \mathcal{X}_t -measurable random vectors of the appropriate dimension. Using the law of iterated expectations and the representation (11) for

¹⁴Orders of minimization and maximization can be exchanged and alternative timing protocols support the same equilibrium outcomes. See Basar and Bernhard (1995) and Hansen, Sargent, Turmuhambetova, and Williams (2004) for complete discussions of these issues. Hansen, Sargent, Turmuhambetova, and Williams (2004) use a continuous-time specification with a Brownian motion information structure.

M_∞ that prevails when $\rho = 1$, we can represent this game in the alternative form

$$-\frac{1}{2} \max_{\{a_t \in \mathcal{A}_t^*\}} \min_{\{m_{t+1} \in \mathcal{M}_{t+1}\}} E \left[\sum_{t=0}^{\infty} M_t \left(\beta^t \begin{bmatrix} a_t' & x_t' \end{bmatrix} \begin{bmatrix} Q & P \\ P' & R \end{bmatrix} \begin{bmatrix} a_t \\ x_t \end{bmatrix} + \theta m_{t+1} \log m_{t+1} \right) | \mathcal{X}_0 \right]$$

subject to:

$$\begin{aligned} x_{t+1} &= Ax_t + Ba_t + Cw_{t+1} \\ M_{t+1} &= m_{t+1}M_t. \end{aligned}$$

We set the initial value of M_0 at unity, which eliminates the term $E\theta M_0 \log M_0 | \mathcal{X}_0$. In this specification, $\{M_t : t \geq 0\}$ is a nonnegative martingale adapted to $\{\mathcal{X}_t : t \geq 0\}$, and m_{t+1} is related to v_t in problem (39)–(40) in a way to be described in equation (45).

Problem (43)–(44) simplifies. The quadratic objective makes the worst-case probability for the shocks become normal. The distortion consists of a state dependent mean shift and a state independent shift in the covariance matrix. The value function of the Markov perfect equilibrium scales linearly in M_t and equals

$$-\frac{\beta^t}{2} M_t [(x_t)' \Omega_t x_t + \omega_t],$$

where Ω_t satisfies (42). Randomness contributes the constant term ω_t .

The distortion m_{t+1} is a likelihood ratio that changes the distribution of w_{t+1} from a normal distribution with mean zero and covariance matrix I to a normal distribution with a mean v_t that is given by the second equation of (41) and a covariance matrix Υ_t , where $\Upsilon_t^{-1} = I - \frac{1}{\theta} C' \Omega_{t+1} C \beta^{t+1}$. The distortion m_{t+1} equals

$$m_{t+1} = \exp \left[-\frac{1}{2} (w_{t+1} - v_t)' (\Upsilon_t)^{-1} (w_{t+1} - v_t) + \frac{1}{2} w_{t+1} \cdot w_{t+1} - \frac{1}{2} \log \det \Upsilon_t \right],$$

A simple calculation shows that

$$E(m_{t+1} \log m_{t+1} | \mathcal{X}_t) = \frac{1}{2} [|v_t|^2 + \text{trace}(I - \Upsilon_t^{-1}) - \log \det \Upsilon_t] \quad (45)$$

where the component terms $\frac{1}{2}|v_t|^2$ and $\text{trace}(I - \Upsilon_t^{-1}) - \log \det \Upsilon_t$ are both nonnegative.

A form of certainty equivalence prevails in the sense that the solutions for a_t, v_t as functions of x_t equal those given by (41). The left side of (45) converges to zero as t gets large under the distorted distribution. Thus, all components of the right side of (45) converge to zero at rates sufficiently fast to guarantee convergence of M_t to M_∞ . Therefore, when $\rho = 1$, the disagreement between the approximating and worst case models vanishes for tail events.

7.4 Hidden states

We complete the solution of problem (37)–(38) by concealing some of the states. The admissible control set \mathcal{A}_t now contains random vectors that are \mathcal{S}_t measurable. We exploit the linear-quadratic structure to show that the ‘temporary approximating model’ $\check{P}r_k$ for z_k from section 6.4 is a Gaussian distribution with mean \check{z}_k and conditional covariance $\check{\Sigma}_k$, and

we give recursions for computing $\check{z}_k, \check{\Sigma}_k$. In appendix B we show how to match our recursions with different ones that Basar and Bernhard (1995) derived in a nonstochastic setting.¹⁵ The appendix B algorithm allows us to interpret $\check{q}_t(z)$ as a distorted Kalman filter that makes operational the recursive characterization of the commitment solution set forward in section 6.4. In our recursive implementation of the commitment solution, the date k minimizing player treats these distorted densities as part of the date k approximating model.

7.4.1 Ordinary Kalman filter for the hidden state

Our analysis starts with the ordinary Kalman filter taking the approximating model as true. Write the hidden state evolution equation as:

$$z_{t+1} = A_{21}y_t + A_{22}z_t + B_2a_t + C_2w_{t+1},$$

and the signal equation as:

$$s_{t+1} = D_1y_t + D_2z_t + Gw_{t+1}.$$

The ordinary Kalman filter for predicting z_t using the original benchmark model is:

$$\bar{z}_{t+1} = A_{21}y_t + A_{22}\bar{z}_t + B_2a_t + K_2(\Delta_t)[s_{t+1} - D_1y_t - D_2\bar{z}_t]$$

where

$$K_2(\Delta) = (A_{22}\Delta D_2' + C_2G')(D_2\Delta D_2' + GG')^{-1}.$$

The covariance matrix is updated using:

$$\Delta_{t+1} = A_{22}\Delta_t A_{22}' + C_2C_2' - (A_{22}\Delta_t D_2' + C_2G')(D_2\Delta_t D_2' + GG')^{-1}(A_{22}\Delta_t D_2' + C_2G)'$$

Thus, under the approximating model, $z_{t+1} \sim \mathcal{N}(\bar{z}_{t+1}, \Delta_{t+1})$.

7.4.2 Altered distribution

Partition P and R as:

$$P = \begin{bmatrix} P_1 & P_2 \end{bmatrix}$$

and

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}.$$

Then write:

$$\begin{bmatrix} a' & x' \end{bmatrix} \begin{bmatrix} Q & P \\ P' & R \end{bmatrix} \begin{bmatrix} a \\ x \end{bmatrix} = \begin{bmatrix} a' & y' & z' \end{bmatrix} \begin{bmatrix} Q & P_1 & P_2 \\ P_1' & R_{11} & R_{12} \\ P_2' & R_{12}' & R_{22} \end{bmatrix} \begin{bmatrix} a \\ y \\ z \end{bmatrix}.$$

For this objective, the date t ingredient corresponding to (33) for constructing the sequence of distorted benchmark models is:

$$\check{m}_t = \exp \left(-\frac{\beta^t}{2\theta} \begin{bmatrix} a_t' & y_t' & z_t' \end{bmatrix} \begin{bmatrix} Q & P_1 & P_2 \\ P_1' & R_{11} & R_{12} \\ P_2' & R_{12}' & R_{22} \end{bmatrix} \begin{bmatrix} a_t \\ y_t \\ z_t \end{bmatrix} \right).$$

¹⁵Also see Whittle (1990).

The density under the original benchmark model (i.e., the approximating model at time 0) is normal with mean \bar{z}_t and covariance matrix Δ_t . By the usual complete the square argument, the \check{m}_t distortion reduces the precision matrix (inverse of the covariance matrix) by

$$(\tilde{\Delta}_t)^{-1} = (\Delta_t)^{-1} - \frac{\beta^t}{\theta} R_{22} \quad (46)$$

and alters the mean by

$$\begin{aligned} \tilde{z}_t &= \left[(\Delta_t)^{-1} - \frac{\beta^t}{\theta} R_{22} \right]^{-1} \left[\frac{1}{\theta} P_2' a_t + \frac{1}{\theta} R_{12}' y_t + (\Delta_t)^{-1} \bar{z}_t \right] \\ &= \bar{z}_t + \frac{\beta^t}{\theta} \tilde{\Delta}_t (P_2' a_t + R_{12}' y_t + R_{22} \bar{z}_t). \end{aligned} \quad (47)$$

When hidden states do not enter the objective function, $R_{12} = 0$, $R_{22} = 0$, and $P_2 = 0$, which according to (46) and (47) implies that there is no distortion. Otherwise, there is a distortion that depends on objective function parameters, controls, and predicted states.

These calculations justify the following algorithm to update the distorted mean and variance. Take the date t distorted mean and covariance matrix and update both using the ordinary Kalman filter, then distort these updates in accordance to formulas (46) and (47). Formally, this results in the *two step* recursion. The outcome of the first step is a conditional mean \check{z}_{t+1} and a conditional covariance matrix $\check{\Delta}_{t+1}$, which are convenient sufficient statistics for the conditional normal density \check{q}_{t+1} . The density \check{q}_{t+1} serves as a benchmark for the date $t + 1$ conditional problem. The preceding results tell us that the ‘temporary benchmark density’ \check{q}_{t+1} can be computed in two steps. The first step is to take the distorted density \check{q}_t and to compute $(\tilde{z}_t, \tilde{\Delta}_t)$ from:

$$\begin{aligned} \tilde{z}_t &= \frac{\beta^t}{\theta} \tilde{\Delta}_t (P_2' a_t + R_{12}' y_t + R_{22} \check{z}_t) + \check{z}_t \\ (\tilde{\Delta}_t)^{-1} &= (\check{\Delta}_t)^{-1} - \frac{\beta^t}{\theta} R_{22}. \end{aligned}$$

The second step starts from $\tilde{z}_t, \tilde{\Delta}_t$ and gets $\check{q}_{t+1}(z)$ by taking one iteration on the ordinary Kalman filter:

$$\begin{aligned} \check{z}_{t+1} &= A_{21} y_t + A_{22} \tilde{z}_t + B_2 a_t + K_2 (\tilde{\Delta}_t) (s_{t+1} - D_1 y_t - D_2 \tilde{z}_t) \\ \check{\Delta}_{t+1} &= A_{22} \tilde{\Delta}_t A_{22}' + C_2 C_2' \\ &\quad - (A_{22} \tilde{\Delta}_t D_2' + C_2 G') (D_2 \tilde{\Delta}_t D_2' + G G')^{-1} (A_{22} \tilde{\Delta}_t D_2' + C_2 G')' \end{aligned}$$

with $(\tilde{z}_t, \tilde{\Delta}_t)$ as state variables.

Appendix B shows that combining these two recursions into one attains versions of recursions that Basar and Bernhard (1995) derived in a nonstochastic setting.

7.5 Worst-case certainty equivalence

We can use the solution to the full information model along with \check{z}_k to represent the robust control law when some states are hidden. Recall the full information value function:

$$-\frac{1}{\theta} \beta^k M_k [(x_k)' \Omega_k x_k + \omega_k].$$

Our use of this value function is an application of a modified version of certainty equivalence. The full information distortion m_{t+1} in this context pretends that the action a_t can condition on information in \mathcal{X}_t instead of \mathcal{S}_t because of certainty equivalence. This trick simplifies computation because it does not involve filtering. After this problem has been solved, the solution of the limited information problem can be obtained by replacing the known state from the full information decision rule with an estimate of the state adjusted for robustness.

At date k , factor $M_k = h_k G_k$ and solve the conditional problem:

$$\min_{h_k \in \mathcal{H}_k} \check{E} \left(\beta^k [(x_k)' \Omega_k x_k + \omega_k] + \theta h_k \log h_k \mid \mathcal{S}_k \right) \quad (48)$$

Under both the original and the $(\check{\cdot})$ probability model, the conditional covariance matrix of x_k conditioned on \mathcal{S}_k is singular because y_k is known at date k ; z_k is normal with mean \check{z}_k and covariance matrix $\check{\Delta}_k$. The minimizer of (48) is:

$$h_k = \frac{\exp \left[-\frac{\beta^k}{2\theta} (x_k)' \Omega_k x_k \right]}{\check{E} \left(\exp \left[-\frac{\beta^k}{2\theta} (x_k)' \Omega_k x_k \right] \mid \mathcal{S}_k \right)}.$$

It implies that z_k has a distorted distribution with mean:

$$\hat{z}_k = \check{z}_k + \left[(\Delta_k)^{-1} - \frac{\beta^k}{\theta} \begin{bmatrix} 0 & I \end{bmatrix} \Omega_k \begin{bmatrix} 0 \\ I \end{bmatrix} \right]^{-1} \frac{\beta^k}{\theta} \Omega_k \begin{bmatrix} y_k \\ \check{z}_k \end{bmatrix}.$$

The robust control obeys a *modified certainty equivalence principle* because it is given by the robust control law (41) for a_k , which we computed under full information, but now to be evaluated at the distorted state estimate \hat{z}_k . The a_k decision is

$$a_k = - \begin{bmatrix} I & 0 \end{bmatrix} \left[\beta \left(\tilde{Q}_k + \beta \tilde{B}' \Omega_{k+1} \tilde{B} \right)^{-1} \tilde{B}' \Omega_{k+1} \tilde{A} + (Q_t)^{-1} \tilde{P} \right] \begin{bmatrix} y_k \\ \hat{z}_k \end{bmatrix}. \quad (49)$$

As emphasized by Whittle (1990), the decision rule (49) has forward looking components that come from ‘control’ and backward looking components that come from ‘filtering under commitment’. Under commitment, the sufficient statistic \check{z}_k used as a benchmark for distorting in state estimation is backward looking. When hidden state variables enter the one-period utility function, \check{z}_k can deviate from the state estimate obtained by direct application of the Kalman filter. The forward-looking component comes from the control component of the problem through the matrices Ω_{k+1} and Ω_k in (42). Both components are combined to produce a distorted estimate of the hidden state \hat{z}_k and the robust action a_k .

7.6 Examples

We consider two examples each of which has a special structure.

The first example is a pure estimation problem.

Example 7.1. *Suppose that the state cannot be influenced by the decision-maker and that the objective is to estimate $-Px_t$. Let the control be simply a robust estimate of $-Px_t$. Set $Q = I$ and $R = P'P$. For this problem $\Omega_t = 0$ because $a = -Px$ sets the full information objective to zero. The solution to the estimation problem is $a_t = -P\hat{x}_t$.*

The second example supposes that the objective does not depend on the hidden state.

Example 7.2. *Suppose that $P_2 = 0$, $R_{12} = 0$, and $R_{22} = 0$. There is another way to solve the robust control problem that first solves the filtering problem and then computes an ordinary robust control for the reduced information configuration by applying the approach described in section 6.6 in conjunction with certainty equivalence.*

Write the solution to the filtering problem as:

$$\bar{z}_{t+1} = A_{21}y_t + A_{22}\bar{z}_t + B_2a_t + K_2(\Delta_t)\bar{w}_{t+1}$$

where

$$\bar{w}_{t+1} = D_2(z_t - \bar{z}_t) + Gw_{t+1}.$$

The constructed disturbance is normal with means zero and covariance matrix:

$$D\Sigma_t D' + GG'$$

conditioned on \mathcal{S}_t . Instead of distorting the joint distribution (w_{t+1}, x_t) , we can distort the distribution of \tilde{w}_{t+1} conditioned on \mathcal{S}_t . This amounts to focusing directly on the g_{t+1} . For computational purposes it suffices to focus on a mean shift distortion \bar{v}_t with entropy penalty:

$$\theta \bar{v}_t' (D\Sigma_t D' + GG')^{-1} \bar{v}_t$$

where the mean shift is now restricted to be a function of the signal history. Despite there also being a bigger conditional covariance, certainty equivalence allows us to compute the mean distortion by solving a deterministic two-player game. As in the robustness problem with full information, discounting causes the impact of discounting to wear off over time.

8 Concluding remarks

Inspired by the pair (31)-(32) that give a recursive representation of the commitment game in the special case that $\rho = 1$, Hansen and Sargent (2005) define operators \mathbb{T}^1 and \mathbb{T}^2 that correspond to \mathbb{R}^1 and \mathbb{R}^2 , respectively, and that apply when the state can be taken to be (y_t, q_t) . They analyze a zero sum game associated with the following recursions

$$W(x, q) = \mathbb{T}^1 [U(x, a) + \beta W(x^*, q^*) | \theta_1] (x, q, a), \quad (50)$$

after choosing an optimal action by solving

$$\max_a \mathbb{T}^2 \{ (\mathbb{T}^1 [U(x, a) + \beta W(y^*, q^*, z^*) | \theta_1] | \theta_2) \} (y, q, a) \quad (51)$$

for $\theta_1 \neq \theta_2$. When $\theta_1 = \theta_2 = \theta$, (50)-(51) alters the setup in this paper by withdrawing commitment from the two players in a zero-sum game, forcing them to choose sequentially, and replacing $\frac{\theta}{\beta^t}$ in (31)-(32) with use θ , thereby discounting increments to entropy at the rate β . By doing this, the formulation in Hansen and Sargent (2005) arrests the decaying of concerns about robustness that leads to the time varying decision rules that emerge in the commitment setup of this paper when $\beta < 1$.

For a finite θ_1 , the operator \mathbb{T}^1 captures the decision maker's fear that the state and signal dynamics conditioned on all components of the state, both hidden and observed, are misspecified, and for a finite θ_2 , the operator \mathbb{T}^2 captures the decision maker's fear that the distribution of the hidden state conditioned on the history of signals is misspecified. Setting $\theta_1 = \theta_2$ focuses the decision maker's attention equally on misspecifications of the state and signal dynamics conditioned on all components of the state, on the one hand, and the distribution of the hidden state conditioned on the history of signals, on the other.

Hansen and Sargent (2004b) advocate separating the values of θ_1 and θ_2 in the operators \mathbb{T}^1 and \mathbb{T}^2 and describe how alternative choices of (θ_1, θ_2) focus the decision maker's distrust on different aspects of his specification. This flexibility will be especially useful in continuous time versions of our model.

A Martingale Convergence

We begin with some useful inequalities.

Lemma A.1. *The function:*

$$\phi(x) = x \log x - x + 1 - (x^{1/2} - 1)^2$$

is nonnegative on $x \geq 0$ where we define $\phi(0)$ as the continuous extension.

Proof. Note that

$$\begin{aligned}\phi'(x) &= \log x + x^{-1/2} - 1 \\ \phi''(x) &= x^{-1} - \frac{1}{2}x^{-3/2}.\end{aligned}$$

The function ϕ is zero at $x = 1$ and $x = 0$. The second derivative has single zero at $x^* = 1/4$. The second derivative is positive to the right and negative to the left of this point. Thus ϕ is convex on the interval $x > x^*$ and concave on the interval $(0, x^*)$. The first derivative ϕ' is zero at $x = 1$, which is necessarily a global minimum over the region $(x^*, +\infty)$. The function ϕ is concave over $(0, x^*)$ and must be positive because $\phi(x^*) > 0$ and $\phi(0) = 0$. \square

We use this inequality to relate entropy to an L^1 approximation.

Lemma A.2. *Consider a nonnegative random variable M with unit expectation. Then*

$$EM \log M \geq E[(M^{1/2} - 1)^2] \geq \frac{1}{4} (E|M - 1|)^2.$$

Proof. First, note that

$$0 \leq E\phi(M) = EM \log M - E[(M^{1/2} - 1)^2],$$

which proves the first inequality. Second, note that the Cauchy-Schwarz and Triangle Inequalities imply that

$$\begin{aligned}E(|M - 1|) &= E [|(M^{1/2} - 1)(M^{1/2} + 1)|] \\ &\leq [E(M^{1/2} - 1)^2]^{1/2} [E(M^{1/2} + 1)^2]^{1/2} \\ &\leq [E(M^{1/2} - 1)^2]^{1/2} [(EM)^{1/2} + 1] \\ &= 2E [(M^{1/2} - 1)^2]^{1/2},\end{aligned}$$

which proves the second inequality. \square

Remark A.3. *The distance*

$$H(M, \tilde{M}) = \left[\frac{1}{2} E |M^{1/2} - \tilde{M}^{1/2}|^2 \right]^{1/2}$$

is referred to as the Hellinger distance between the “derivatives” M and \tilde{M} . It is less than or equal to unity, since $EM = E\tilde{M} = 1$. The second inequality of Lemma A.2 is implied by the well known link between the Hellinger distance and the L^1 norm:¹⁶

$$H(M, \tilde{M})^2 \leq \frac{1}{2}E|M - \tilde{M}| \leq H(M, \tilde{M})\sqrt{2 - H^2(M, \tilde{M})}$$

Replacing the square root term by its upper bound $\sqrt{2}$ gives the second inequality with $\tilde{M} = 1$.

Let $\mathcal{X}_\infty \doteq \bigvee_t \mathcal{X}_t$.

Proposition A.4. Consider a nonnegative martingale $\{M_t : t \geq 0\}$ with unit expectation conditioned on \mathcal{X}_0 and adapted to $\{\mathcal{X}_t : t \geq 0\}$. Suppose that

$$\sum_{t=0}^{\infty} E[M_t E(m_{t+1} \log m_{t+1} | \mathcal{X}_t) | \mathcal{X}_0] < \infty.$$

Then there exists an \mathcal{X}_∞ measurable nonnegative random variable M_∞ such that

$$\lim_{t \rightarrow \infty} E(|M_t - M_\infty| | \mathcal{X}_0) = 0.$$

Proof. Construct

$$M_{t+j}^t = \prod_{k=1}^j m_{t+k}$$

and write:

$$|M_{t+j} - M_t| = M_t |M_{t+j}^t - 1|.$$

Given $\epsilon > 0$ choose t^* such that

$$\sum_{k=1}^{\infty} E[M_{t^*+k-1} E(m_{t^*+k} \log m_{t^*+k} | \mathcal{X}_{t^*}) | \mathcal{X}_0] \leq \frac{1}{4}\epsilon^2,$$

and let $t \geq t^*$. Construct a distorted probability with expectation operator \tilde{E} using the \mathcal{X}_t measurable random variable M_t . Note that the implied probability conditioned on \mathcal{X}_t is not distorted and that

$$\begin{aligned} \tilde{E}(M_{t+j}^t \log M_{t+j}^t | \mathcal{X}_0) &= \tilde{E}\left(M_{t+j}^t \sum_{k=1}^j \log m_{t+k} | \mathcal{X}_0\right) \\ &= \sum_{k=1}^j \tilde{E}(M_t^{t+k} \log m_{t+k} | \mathcal{X}_0) \\ &= \sum_{k=1}^j \tilde{E}[M_t^{t+k-1} E(m_{t+k} \log m_{t+k} | \mathcal{X}_{t+k-1}) | \mathcal{X}_0] \\ &= \sum_{k=1}^j E[M_{t+k-1} E(m_{t+k} \log m_{t+k} | \mathcal{X}_{t+k-1}) | \mathcal{X}_0] \end{aligned}$$

¹⁶See LeCam and Yang (1989), page 25, for more about these inequalities.

$$\leq \frac{1}{4}\epsilon^2$$

for all $j \geq 1$. Moreover,

$$E(|M_{t+j} - M_t| | \mathcal{X}_0) = \tilde{E}(|M_t^{t+j} - 1| | \mathcal{X}_0)$$

From Lemma A.2,

$$\begin{aligned} \tilde{E}(|M_{t+j}^t - 1| | \mathcal{X}_0) &\leq 2 \left[\tilde{E}(M_{t+j}^t \log M_{t+j}^t | \mathcal{X}_0) \right]^{1/2} \\ &\leq \epsilon. \end{aligned}$$

Thus

$$E(|M_{t+j} - M_t| | \mathcal{X}_0) \leq \epsilon$$

for all $j \geq 1$.

Since ϵ was chosen arbitrarily, the martingale sequence $\{M_t : t \geq 0\}$ is conditionally Cauchy in the conditional (on \mathcal{X}_0) L^1 space constructed using the underlying probability measure. This space is conditionally complete and hence this sequence has a limit point M_∞ that is \mathcal{X}_∞ measurable.¹⁷

□

Corollary A.5. *Under the assumptions of Proposition A.4,*

$$M_t = E(M_\infty | \mathcal{X}_t).$$

This follows because for any bounded random variable W_t that is \mathcal{X}_t measurable,

$$\begin{aligned} E(M_\infty W_t | \mathcal{X}_0) &= \lim_{j \rightarrow \infty} E(M_{t+j} W_t | \mathcal{X}_0) \\ &= \lim_{j \rightarrow \infty} E(M_t W_t | \mathcal{X}_0) \\ &= E(M_t W_t | \mathcal{X}_0). \end{aligned}$$

Corollary A.6. *In the statement and conclusion of Theorem A.4, the sigma algebra \mathcal{X}_0 can be replaced by a smaller sigma algebra \mathcal{S}_0 .*

Remark A.7. *Important components of this analysis are anticipated in Kabanov, Lipcer, and Sirjaev (1977). For instance, the criterion of Proposition A.4 implies the sufficient condition for absolute continuity of the distorted probability measure with respect to the original probability measure on \mathcal{X}_∞ given in Corollary 3 of Kabanov, Lipcer, and Sirjaev (1977). Instead of conditional L^1 convergence, Kabanov, Lipcer, and Sirjaev (1977) establish almost sure convergence of the martingale under the distorted probability distribution as a necessary and sufficient condition for absolute continuity.*

¹⁷See Hansen and Richard (1987) for a development of conditional notions of convergence and completeness.

B Robust Control Recursions

In this appendix, we show the connection between the recursive formulas under commitment reported in section 7 and related formulas of Basar and Bernhard (1995). We accomplish this by deriving a different representation of the Kalman filtering updating equation that is appropriate when the conditional covariance matrix Δ_t is nonsingular.

First, transform the date $t + 1$ state to remove correlation with the signal conditioned on x_t :

$$x_{t+1} - CG'(GG')^{-1}s_{t+1} = [A - CG'(GG')^{-1}D]x_t + Ba_t + (C - CG'(GG')^{-1}G)w_{t+1}.$$

In particular, the equation for the second partition is

$$\begin{aligned} z_{t+1} - C_2G'(GG')^{-1}s_{t+1} &= A_{21}y_t + A_{22}z_t - C_2G'(GG')^{-1}(D_1y_t + D_2z_t) \\ &\quad + B_2a_t + [C_2 - C_2G'(GG')^{-1}G]w_{t+1}. \end{aligned}$$

Thus

$$E(z_{t+1}|\mathcal{S}_{t+1}) = C_2G'(GG')^{-1}[s_{t+1} - D_1y_t - D_2E(z_t|\mathcal{S}_{t+1})] + A_{21}y_t + A_{22}E(z_t|\mathcal{S}_{t+1}) + B_2a_t.$$

Next update the conditional precision matrix for z_t given \mathcal{S}_{t+1} :

$$(\Delta_t)^{-1} + D_2'(GG')^{-1}D_2.$$

Recall that $(\Delta_t)^{-1}$ is the precision matrix conditioned on \mathcal{S}_t . Incorporating the information in the signal s_{t+1} increases the precision by $D_2'(GG')^{-1}D_2$. The corresponding prediction formula is:

$$E(z_t|\mathcal{S}_{t+1}) = \bar{z}_t + [(\Delta_t)^{-1} + D_2'(GG')^{-1}D_2]^{-1}D_2'(GG')^{-1}(s_{t+1} - D_1y_t - D_2\bar{z}_t)$$

where $\bar{z}_t \doteq E(z_t|\mathcal{S}_t)$.

Let

$$\begin{aligned} \check{A} &\doteq A_{22} - C_2G'(GG')^{-1}D_2 \\ \check{N} &\doteq C_2C_2' - C_2G'(GG')^{-1}GC_2' \end{aligned}$$

Combining the previous calculations, we obtain the covariance recursion:

$$\Delta_{t+1} = \check{A}[(\Delta_t)^{-1} + D_2'(GG')^{-1}D_2]\check{A}' + \check{N}, \quad (52)$$

and the mean recursion:

$$\begin{aligned} \bar{z}_{t+1} &= A_{21}y_t + \check{A}\bar{z}_t + B_2a_t \\ &\quad + \check{A}[(\Delta_t)^{-1} + D_2'(GG')^{-1}D_2]^{-1}D_2'(GG')^{-1}(s_{t+1} - D_1y_t - D_2\bar{z}_t) \end{aligned} \quad (53)$$

Updating equations (52) and (53) give representations of the Kalman filtering equations that are valid when the covariance matrix Δ_t is nonsingular.

It follows from the covariance updating equation (52) that

$$\check{\Delta}_{t+1} = \check{A}[(\check{\Delta}_t)^{-1} + D_2'(GG')^{-1}D_2]\check{A}' + \check{N}$$

$$= \check{A} \left[(\check{\Delta}_t)^{-1} - \frac{\beta^t}{\theta} R_{22} + D_2'(GG')^{-1}D_2 \right] \check{A}' + \check{N} \quad (54)$$

where we have used the relation:

$$(\tilde{\Delta}_t)^{-1} = (\check{\Delta}_t)^{-1} - \frac{\beta^t}{\theta} R_{22}.$$

Similarly, it follows from the updating equation (53) for the mean that

$$\check{z}_{t+1} = A_{21}y_t + \check{A}\check{z}_t + B_2a_t + \check{A} \left[(\tilde{\Delta}_t)^{-1} + D_2'(GG')^{-1}D_2 \right]^{-1} D_2'(GG')^{-1}(s_{t+1} - D_1y_t - D_2\check{z}_t).$$

Since

$$\tilde{z}_t = \check{z}_t + \frac{\beta^t}{\theta} \tilde{\Delta}_t (P_2'a_t + R_{12}'y_t + R_{22}\check{z}_t),$$

$$\begin{aligned} \check{z}_{t+1} &= A_{21}y_t + \check{A}\check{z}_t + B_2a_t \\ &+ \check{A} \left[(\check{\Delta}_t)^{-1} - \frac{\beta^t}{\theta} R_{22} + D_2'(GG')^{-1}D_2 \right]^{-1} D_2'(GG')^{-1}(s_{t+1} - D_1y_t - D_2\check{z}_t) \\ &+ \frac{\beta^t}{\theta} \check{A} \left[(\check{\Delta}_t)^{-1} - \frac{\beta^t}{\theta} R_{22} + D_2'(GG')^{-1}D_2 \right]^{-1} (P_2'a_t + R_{12}'y_t + R_{22}\check{z}_t). \end{aligned} \quad (55)$$

Updating equations (54) and (55) provide the formulas for producing the reference density sequence $\{\check{q}_t : t \geq 0\}$ recursively. These equations agree with equations (6.61) and (6.62) in Basar and Bernhard (1995) derived for deterministic robust control problems. As expected, when $P_2 = 0$, $R_{12} = 0$ and $R_{22} = 0$ these recursions agree with the updating equations for the original equations (52) and (53) for the Kalman filter.

References

- Anderson, E., L. Hansen, and T. Sargent (2003). A quartet of semigroups for model specification, robustness, prices of risk, and model detection. *Journal of the European Economic Association* 1(1), 68–123.
- Basar, T. and P. Bernhard (1995). *H^∞ -Optimal Control and Related Minimax Design Problems* (second ed.). Birkhauser.
- Bergemann, D. and J. Valimaki (1996). Learning and strategic pricing. *Econometrica* 64, 1125–1149.
- Blackwell, D. and M. A. Girshick (1954). *Theory of Games and Statistical Decisions*. New York: Wiley Publications in Statistics.
- Brock, W. A. and L. J. Mirman (1972). Optimal economic growth and uncertainty: The discounted case. *Journal of Economic Theory* 4(3), 479–513.
- Chamberlain, G. (2000). Econometric applications of maxmin expected utility theory. *Journal of Applied Econometrics* 15, 625–644.
- Cogley, T., R. Colacito, and T. Sargent (2005). Benefits from u.s. monetary policy experimentation in the days of samuelson and solow and lucas. unpublished.
- Dupuis, P. and R. S. Ellis (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley Series in Probability and Statistics. New York: John Wiley and Sons.
- Elliott, R. J., L. Aggoun, and J. B. Moore (1995). *Hidden Markov Models: Estimation and Control*. New York: Springer-Verlag.
- Epstein, L. and M. Schneider (2003, November). Recursive multiple priors. *Journal of Economic Theory* 113(1), 1–31.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (1995). *Bayesian Data Analysis*. Chapman and Hall.
- Hamilton, J. D. (1989, March). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57(2), 357–384.
- Hansen, L. P. and S. F. Richard (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica* 55, 587–614.
- Hansen, L. P. and T. Sargent (1995, May). Discounted linear exponential quadratic gaussian control. *IEEE Transactions on Automatic Control* 40(5), 968–971.
- Hansen, L. P. and T. J. Sargent (2004a). Misspecification in recursive macroeconomic theory. Princeton University Press, forthcoming.
- Hansen, L. P. and T. J. Sargent (2004b). Recursive robust decisions with hidden states. unpublished.
- Hansen, L. P. and T. J. Sargent (2005). Recursive robust decisions with hidden states. unpublished.

- Hansen, L. P., T. J. Sargent, G. A. Turmuhambetova, and N. Williams (2004). Robust control, min-max expected utility, and model misspecification. manuscript, University of Chicago and New York University.
- Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy* 87(5), 972–990.
- Jovanovic, B. (1982, May). Selection and the evolution of industry. *Econometrica* 50(3), 649–670.
- Jovanovic, B. and Y. Nyarko (1995). The transfer of human capital. *Journal of Economic Dynamics and Control* 19, 1033–1064.
- Jovanovic, B. and Y. Nyarko (1996, November). Learning by doing and the choice of technology. *Econometrica* 64(6), 1299–1310.
- Kabanov, J. U., R. S. Lipcer, and A. N. Sirjaev (1977). Absolute continuity and singularity of measures. *Mathematics of the USSR-Sbornik* 33, 203–221.
- Knox, T. (2003). Foundations for learning how to invest when returns are uncertain. unpublished.
- LeCam, L. and G. L. Yang (1989). *Asymptotics in Statistics*. New York: Springer-Verlag.
- Maccheroni, F., M. Marinacci, and A. Rustichini (2004). Variational representation of preferences under ambiguity. unpublished.
- Merton, R. C. (1975, July). The asymptotic theory of growth under uncertainty. *Review of Economic Studies* 42(3), 375–393.
- Petersen, I. R., M. R. James, and P. Dupuis (2000). Minimax optimal control of stochastic uncertain systems with relative entropy constraints. *IEEE Transactions on Automatic Control* 45, 398–412.
- Sclove, S. L. (1983, February). Time-series segmentation: A model and a method. *Information Sciences* 29(1), 7–25.
- Wang, T. (2001). Two classes of multi-prior preferences. unpublished.
- Whittle, P. (1981). Risk sensitive linear quadratic gaussian control. *Advances in Applied Probability* 13, 764–777.
- Whittle, P. (1990). *Risk-Sensitive Optimal Control*. New York: John Wiley & Sons.
- Wonham, W. J. (1964). Some applications of stochastic differential equations to optimal nonlinear filtering. *Siam Journal of Control* 2(3), 347–368.