Robust Control and Model Misspecification

Lars Peter Hansen Thomas J. Sargent
Gauhar A. Turmuhambetova Noah Williams*

March 25, 2004

Abstract

A decision maker fears that the data are generated by a statistical perturbation of an approximating model that is specified either as a controlled diffusion or as a controlled measure over continuous functions of time. For each of these mathematical formulations of the approximating model, the perturbations to his approximating model feared by the decision maker are constrained by relative entropy. Several two-player zero-sum games that yield robust decision rules are related to one another and to the max-min expected utility theory of Gilboa and Schmeidler (1989). Alternative sequential and non-sequential versions of robust control theory yield identical robust decision rules that are dynamically consistent in a useful sense.

1 Introduction

A *decision maker* consists of (i) a utility function that is maximized subject to (ii) a model. Classical decision theory and control theory both assume that a decision maker has complete confidence in his model. Robust control theory presents alternative formulations of a decision maker who doubts his model. To capture the idea that the decision maker views his model as an approximation, these formulations alter items (i) and (ii) by (1) surrounding the decision maker's approximating model with a cloud of other models that are difficult to distinguish from it with finite data, and (2) adding a malevolent second agent. The malevolent agent promotes robustness by causing the decision maker to explore the fragility of candidate decision rules to departures of the data from the approximating model. Finding a rule that is robust to model misspecification entails computing bounds on a rule's performance. The minimizing agent helps the maximizing agent construct bounds.

Robust control theory uses alternative mathematical formalisms. While all of them have versions of items (1) and (2), they differ in many important mathematical details including

*We thank Fernando Alvarez, Gary Chamberlain, Ivar Ekeland, Christian Gourieroux, Peter Klibanoff, Michael Allen Rierson, Aldo Rustichini, Jose Scheinkman, Christopher Sims, Nizar Touzi, and especially Costis Skiadas for valuable comments on an earlier draft. Sherwin Rosen encouraged us to write this paper.

the probability spaces on which they are defined; their ways of representing alternative models; the restrictions that they impose on the set of alternative models; and the timing protocols that they impose on the maximizing and minimizing decision makers. Nevertheless, common outcomes and representations emerge from these alternative formulations. That fact allows us to express the same economic content in terms of different mathematical formulations and thereby enables diverse interpretations. Equivalent concerns about model misspecification can be represented by either (a) altering the decision maker's preferences to enhance risk-sensitivity, or (b) leaving his preferences alone but slanting his expectations relative to his approximating model in a particular context-specific way, or (c) adding a set of perturbed models and a malevolent agent. This paper exhibits these unifying connections and stresses how they can be exploited in applications.

Robust control theory shares with both the Bayesian paradigm and the rational expectations model the feature that the decision maker brings to the table only one fully specified model. In robust control theory it is called either his reference model or his approximating model. Although the decision maker does not explicitly specify other models, he evaluates a decision rule under a set of incompletely articulated models that are formed by perturbing his approximating model. Robust control theory contributes useful ways of surrounding a single approximating model with a cloud of models. We can regard that set of models as the multiple priors that appear in the max-min expected utility theory of Gilboa and Schmeidler (1989). We give technical conditions that validate this interpretation in a dynamic setting. Some of these technical issues are about how to represent the approximating model and perturbations to it; others involve conditions that make equilibrium outcomes of various two-player zero-sum games become identical.

This paper starts with two alternative ways of representing an approximating model in continuous time – either (1) as a diffusion or (2) as a measure over continuous functions of time that are induced by the diffusion. We consider different ways of perturbing each such representation of the approximating model. These alternative representations of the approximating model and the perturbations to it lead to alternative formulations of robust control problems. In all of our problems, we use a definition of relative entropy (an expected log likelihood ratio) to constrain the gap between the approximating model and a perturbation to it. Our entropy measure must be less than a finite number that we take as a parameter that measures the set of perturbations against which the decision maker seeks robustness. Requiring that entropy be finite restricts the form that model misspecification can take. In particular, finiteness of entropy implies that admissible perturbations of the approximating model must be absolutely continuous with respect to it over finite intervals. For a diffusion, absolute continuity over finite intervals implies that allowable perturbations can alter the drift but not the volatility. Restricting ourselves to perturbations that are absolutely continuous over finite intervals amounts to considering only perturbed models that are in principle statistically difficult to distinguish from the approximating model, an idea exploited by Anderson, Hansen, and Sargent (2003) to calibrate a plausible amount of fear of model misspecification in a study of market prices of risk.

The work of Araujo and Sandroni (1999) and Sandroni (2000) features how absolute continuity of models implies that decision makers' beliefs eventually merge with the model that generates the data. In infinite horizon economies, absolute continuity over finite intervals does not imply absolute continuity. By allowing perturbations that are not absolutely continuous, we arrest merging and thereby create a setting in which a decision maker's fear of model misspecification endures. Perturbations that are absolutely continuous over finite intervals but still not absolutely continuous can be difficult to detect from a continuous record of finite length, though they could be detected from a continuous data record of infinite length. We discuss how this modelling choice interacts with the way that the decision maker discounts the future.

We also consider a variety of technical issues about timing protocols that underlie interconnections among various expressions of robust control theory. A Bellman-Isaacs condition allows us to exchange orders of minimization and maximization and validates several useful results, including the existence of a Bayesian interpretation of a robust decision rule.

Counterparts to many of the issues treated in this paper occur in discrete time robust control theory. Many of these issues surface in nonstochastic versions of the theory, for example, in Basar and Bernhard (1995). The continuous time stochastic setting of this paper allows sharper analytical results in several cases.

1.1 Organization of paper

The substantial quantity and technical nature of interrelated material inspires us to present it in two exposures consisting first of section 2, then of the remaining sections. Section 2 sets aside a variety of complications and compiles our main results by displaying Bellman equations for various games and decision problems and asserting without proof the key relationships among them. The remaining sections lay things out in much greater detail. Section 3 sets the stage by describing both sequential and nonsequential versions of a control problem under a known model. These problems form benchmarks against which to judge subsequent problems in which the decision maker distrusts his model. Section 3 also introduces a risk-sensitive control problem that alters the decision maker's objective function but leaves secure his trust in his model. Section 4 discusses alternative ways of representing fear of model misspecification. Section 5 introduces entropy and its relationship to a concept of absolute continuity over finite intervals, then formulates two nonsequential zero-sum two-player games, called penalty and constraint games, that induce robust decision rules. The games in section 5 are both cast in terms of sets of probability measures. In section 6, we cast counterparts to these games on a fixed probability measure by representing perturbations to an approximating model in terms of martingales defined on a fixed probability space. Section 7 gives a sequential formulation of a penalty game. By taking continuation entropy as an endogenous state variable, section 8 gives a sequential formulation of a constraint game. This formulation sets the stage for our discussion in section 9 of the dynamic consistency issues raised by Epstein and Schneider (2001). Section 10 concludes. Appendix A is a cast of characters that records the objects and concepts that occur throughout the paper. Four additional appendixes deliver proofs.

2 Overview

⟨onethousand⟩ One Bellman equation is worth a thousand words. We can concisely summarize our main results by displaying Bellman equations for various two-player zero-sum continuous time games that are defined in terms of a Markov diffusion with state x and Brownian motion B, together with the value functions for some related static games. Our story is encoded in the different state variables, drifts, and quadratic variation terms that occur in the Bellman equations for the dynamic games. This telegraphic section is intended for readers who can glean everything from Bellman equations. Readers who prefer a more deliberate presentation from the beginning should skip to section 3.

2.1 Dynamic control problems and games

Benchmark control problem:

We take as a benchmark an ordinary control problem with value function

$$J(x_0) = \sup_{c \in C} E\left[\int_0^\infty \exp(-\delta t) U(c_t, x_t) dt\right]$$

where the maximization is subject to $dx_t = \mu(c_t, x_t)dt + \sigma(c_t, x_t)dB_t$ and where x_0 is a given initial condition. The Bellman equation for the benchmark problem is

$$\delta J(\check{x}) = \max_{\check{c} \in \check{C}} U(\check{c}, \check{x}) + \mu(\check{c}, \check{x}) \cdot J_x(\check{x}) + \frac{1}{2}\text{trace}\left[\sigma(\check{c}, \check{x})' J_{xx}(\check{x}) \sigma(\check{c}, \check{x})\right]. \tag{1}$$ intro1

The notation $\check{\cdot}$ is used to denote a potentially realized value of a control or a state. Similarly, $\check{C}$ is the set of admissible values for the control.

In the benchmark problem, the decision maker trusts his model. We want to study comparable problems where the decision maker distrusts his model. Several superficially different devices can be used to promote robustness to misspecification of the diffusion associated with (1). These add either a free parameter $\theta > 0$ or a state variable $\check{r} \geq 0$ or a state vector X and produce recursive problems with one of the following Bellman equations:

Risk sensitive control problem:

$$\begin{aligned}\delta S(\check{x}) &= \max_{\check{c} \in \check{C}} U(\check{c}, \check{x}) + \mu(\check{c}, \check{x}) \cdot S_x(\check{x}) + \frac{1}{2}\text{trace}\left[\sigma(\check{c}, \check{x})' S_{xx}(\check{x}) \sigma(\check{c}, \check{x})\right] \\ &\quad - \frac{1}{2\theta} S_x(\check{x})' \sigma(\check{c}, \check{x}) \sigma(\check{c}, \check{x})' S_x(\check{x}) \end{aligned} \tag{2}$$ intro2

Bellman equation (2) alters the right side of the value function recursion (1) by deducting $\frac{1}{2}\theta$ times the local variation of the continuation value. The optimal decision rule for the risk-sensitive problem (2) is a policy function

$$c_t = \alpha_c(x_t)$$

where the dependence on θ is understood. In the control theory literature, $-1/\theta$ is called the *risk-sensitivity parameter*; in the recursive utility literature, it is called the *variance multiplier*.

Penalty robust control problem:

A two-player zero-sum game has a value function M that satisfies

$$M(\check{x}, \check{z}) = \check{z} V(\check{x}) \tag{3}$$ intro8

where z_t is another state variable that changes the probability distribution and V satisfies:

$$\begin{aligned} \delta V(\check{x}) \;&=\; \max_{\check{c}\in\check{C}} \min_{\check{h}} U(\check{c}, \check{x}) + \frac{\theta}{2}\check{h}\cdot\check{h} + \left[\mu(\check{c}, \check{x}) + \sigma(\check{c}, \check{x})\check{h}\right]\cdot V_x(\check{x}) \\ &\quad + \frac{1}{2}\text{trace}\left[\sigma(\check{c}, \check{x})' V_{xx}(\check{x})\sigma(\check{c}, \check{x})\right]. \end{aligned} \tag{4}$$ intro3

The process $z = \{z_t : t \geq 0\}$ is a martingale with initial condition $z_0 = 1$ and evolution $dz_t = h_t \cdot dB_t$. The minimizing agent in (4) chooses an $\check{h}$ to alter the probability distribution; $\theta > 0$ is a parameter that penalizes the minimizing agent for distorting the drift. Optimizing over h shows that V from (4) solves the same partial differential equation (2).

Constraint robust control problem:

A two-player zero-sum game has a value function $\check{z}K(\check{x}, \check{r})$ where

$$\begin{aligned} \delta K(\check{x}, \check{r}) \;&=\; \max_{\check{c}\in\check{C}} \min_{\check{h},\check{g}} U(\check{c}, \check{x}) + \left[\mu(\check{c}, \check{x}) + \sigma(\check{c}, \check{x})\check{h}\right]\cdot K_x(\check{x}, \check{r}) + \left(\delta\check{r} - \frac{\check{h}\cdot\check{h}}{2}\right)\cdot K_r(\check{x}, \check{r}) \\ &\quad + \frac{1}{2}\text{trace}\left(\begin{bmatrix} \sigma(\check{c}, \check{x})' & \check{g} \end{bmatrix} \begin{bmatrix} K_{xx}(\check{x}, \check{r}) & K_{xr}(\check{x}, \check{r}) \\ K_{rx}(\check{x}, \check{r}) & K_{rr}(\check{x}, \check{r}) \end{bmatrix} \begin{bmatrix} \sigma(\check{c}, \check{x}) \\ \check{g}' \end{bmatrix}\right). \end{aligned} \tag{5}$$ intro4

Equation (5) shares with (4) that the minimizing agent chooses an $\check{h}$ that alters the probability distribution, but unlike (4), there is no penalty parameter θ. Instead, in (5), the minimizing agent's choice of h_t also affects a new state variable r_t that we call *continuation entropy*. The minimizing player also controls another decision variable $\check{g}$ that determines how increments in the continuation value are related to the underlying Brownian motion. The right side of the Bellman equation for the constraint control problem (5) is attained by decision rules

$$c_t = \phi_c(x_t, r_t) \quad h_t = \phi_h(x_t, r_t) \quad g_t = \phi_g(x_t, r_t).$$

We can solve the equation $\frac{\partial}{\partial r}K(x_t, r_t) = -\theta$ to express r_t as a time invariant function of x_t:

$$r_t = \phi_r(x_t).$$

Therefore, along an equilibrium path of game (5), we have $c_t = \phi_c[x_t, \phi_r(x_t)]$, $h_t = \phi_h[x_t, \phi_r(x_t)]$, $g_t = \phi_g[x_t, \phi_r(x_t)]$.

A problem with a Bayesian interpretation:

A single agent optimization problem has a value function $\check{z}W(\check{x}, \check{X})$ where W satisfies:

$$\begin{aligned}
\delta W(\check{x}, \check{X}) &= \max_{\check{c} \in \check{C}} U(\check{c}, \check{x}) + \mu(\check{c}, \check{x}) \cdot W_x(\check{x}, \check{X}) + \mu^*(\check{x}) \cdot W_X(\check{x}, \check{X}) \\
&+ \frac{1}{2} \text{trace} \left(\begin{bmatrix} \sigma(\check{c}, \check{x})' & \sigma^*(\check{X})' \end{bmatrix} \begin{bmatrix} W_{xx}(\check{x}, \check{X}) & W_{xX}(\check{x}, \check{X}) \\ W_{Xx}(\check{x}, \check{X}) & W_{XX}(\check{x}, \check{X}) \end{bmatrix} \begin{bmatrix} \sigma(\check{c}, \check{x}) \\ \sigma^*(\check{X}) \end{bmatrix} \right) \\
&+ \alpha_h(\check{X}) \cdot \sigma(\check{c}, \check{x})' W_x(\check{x}, \check{X}) + \alpha_h(\check{X}) \cdot \sigma^*(\check{X})' W_X(\check{x}, \check{X}) \qquad (6) \; \texttt{intro5}
\end{aligned}$$

where $\mu^*(\check{X}) = \mu[\alpha_c(\check{X}), \check{X}]$ and $\sigma^*(\check{X}) = \sigma[\alpha_c(\check{X}), \check{X}]$. The function $W(\check{x}, \check{X})$ in (6) depends on an additional component of the state vector $\check{X}$ that is comparable in dimension with $\check{x}$ and that is to be initialized from the common value $\check{X}_0 = \check{x}_0 = x_0$. We shall show in appendix E that equation (6) is the Bellman equation for an ordinary (i.e., single agent) control problem with discounted objective:

$$z_0 W(\check{x}, \check{X}) = E \int_0^\infty \exp(-\delta t) z_t U(c_t, x_t) dt$$

and state evolution:

$$\begin{aligned}
dx_t &= \mu(c_t, x_t)dt + \sigma(c_t, x_t)dB_t \\
dz_t &= z_t \alpha_h(X_t) dB_t \\
dX_t &= \mu^*(X_t)dt + \sigma^*(X_t)dB_t
\end{aligned}$$

with $z_0 = 1, x_0 = \check{x}$, and $X_0 = \check{X}$.

This problem alters the benchmark control problem by changing the probabilities assigned to the *shock* process $\{B_t : t \geq 0\}$. It differs from the martingale problem because the process z used to change probabilities does not depend on state variables that are endogenous to the control problem.

In appendix E, we verify that under the optimal c and the prescribed choices of $\mu^*, \sigma^*, \alpha_h$, the 'big X' component of the state vector state variable equals the 'little x' component, provided that $\check{X}_0 = \check{x}_0$. Equation (6) is therefore the Bellman equation for an ordinary control problem that justifies a robust decision rule under a fixed probability model that differs from the approximating model. As the presence of z_t as a preference shock suggests, this problem reinterprets the equilibrium of the two-player zero-sum game portrayed in the martingale problem (3). For a given θ that gets embedded in σ^*, μ^*, the right side of the Bellman equation (6) is attained by $\check{c} = \gamma_c(\check{x}, \check{X})$.

2.1.1 Different ways to attain robustness

Relative to that for the benchmark problem (1), Bellman equations (2), (4), (5), and (6) can all be interpreted as devices that in different ways promote robustness to misspecification of the diffusion. Bellman equations (2) and (6) are for ordinary control problems: only the maximization operator appears on the right side and there is no minimizing player to promote robustness. Instead, problem (2) promotes robustness by altering the maximizing player's preferences to enhance risk sensitivity, while problem (6) promotes robustness by attributing to the maximizing player a belief about the state transition law that is distorted in a particular way relative to his approximating model. Meanwhile, the Bellman equations in (4) and (5) are for two-player zero-sum dynamic games in which a minimizing player promotes robustness.

2.2 Static problems

ubsec:static⟩? We also study two static two-player zero-sum games that are defined in terms of perturbations $q \in Q$ to the measure q^0 over continuous functions of time that is induced by the Brownian motion B in the diffusion for x. We define discounted relative entropy as

$$\tilde{\mathcal{R}}(q) \doteq \delta \int_0^\infty \exp(-\delta t) \left(\int \log \left(\frac{dq_t}{dq_t^0} \right) dq_t \right) dt$$

and use it to restrict the size of perturbations q to q^0. Leaving the dependence on B implicit, we define a utility process $\upsilon_t(c) = U(c_t, x_t)$ and pose the following two problems:

Nonsequential penalty control problem:

$$\tilde{V}(\theta) = \max_{c \in C} \min_{q \in Q} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt + \theta \mathcal{R}^*(q). \tag{7}$$

intro6

Nonsequential constraint control problem:

$$\tilde{K}(\eta) = \max_{c \in C} \min_{q \in Q(\eta)} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt \tag{8}$$

intro7

where $Q(\eta) = \{q \in Q : \tilde{\mathcal{R}}(q) \leq \eta\}$.

Problem (8) fits the max-min expected utility model of Gilboa and Schmeidler (1989), where $Q(\eta)$ is the set of multiple priors in their model. The axiomatic treatment of Gilboa and Schmeidler views this set of priors as an expression of the decision maker's preferences and does not cast them as perturbations of an approximating model.[1] We are free to think of problem (8) as providing a way to use a single approximating model q^0 to generate Gilboa-Schmeidler's set of priors, all those unspecified models that satisfy the restriction on relative entropy $Q(\eta) = \{q \in Q : \tilde{\mathcal{R}}(q) \leq \eta\}$.

The objective functions for these two static optimization problems (7) and (8) are related via the Legendre transform pair:

$$\tilde{V}(\theta) = \min_{\eta \geq 0} \tilde{K}(\eta) + \theta \eta \tag{9}$$

intro9

[1] Similarly, Savage's framework does not purport to describe the process by which the Bayesian decision maker constructs his unique prior.

$$\tilde{K}(\eta) = \max_{\theta \geq 0} \tilde{V}(\theta) - \eta\theta. \qquad (10) \text{ intro10}$$

2.3 Connections

An association between robust control and the framework of Gilboa and Schmeidler (1989) extends beyond problem (8) because the equilibrium value functions and decision rules for all of our problems are intimately related. The recursive counterpart to (9) is:

$$V(\check{x}) = \min_{\check{r} \geq 0} K(\check{x}, \check{r}) + \theta\check{r}$$

with the implied first-order condition

$$\frac{\partial}{\partial r} K(\check{x}, \check{r}) = -\theta$$

that implicitly defines $\check{r}$ as a function of $\check{x}$ for a given θ, which implies that $\check{r}$ is a redundant state variable. The penalty formulation avoids this redundancy.[2]

The static value function $\tilde{V}$ is related to other value functions:

$$\tilde{V}(\theta) = M(\check{x}, \check{z}) = \check{z}V(\check{x}) = W(\check{x}, \check{x}) = S(\check{x}),$$

provided that $\check{x}$ is initialized at x_0 and $\check{z}$ at 1 and that θ is held fixed across the different problems. Though these problems have different decision rules, we shall show that for a fixed θ and comparable initial conditions, they have identical equilibrium outcomes and identical recursive representations of those outcomes.

The following relations prevail across the equilibrium decision rules for our different problems:

$$\alpha_c(\check{x}) = \gamma_c(\check{x}, \check{x}) = \phi_c[\check{x}, \phi_r(\check{x})].$$

2.3.1 Who cares?

Why do we care about the equivalence of these control problems and games? Beyond beauty: some of the problems are easier to solve, but others are easier to interpret.

These problems have emerged from literatures that have approached the problem of decision making in the presence of model misspecification from different angles. The recursive version of the penalty problem (6.4) emerged from a literature on robust control that also considered the risk-sensitive problem (14). The 'static constraint' problem (5.3) is an example of the min-max expected utility theory of Gilboa and Schmeidler (1989) with a particular set of priors. By modifying the set of priors over time, constraint problem (5) states a recursive version of that static constraint problem. The Lagrange multiplier theorem supplies an interpretation of the penalty parameter θ.

[2]There is also a recursive analog to (10) that uses the fact that the function V depends implicitly on θ.

A potentially unattractive feature of multiple priors models for applied work is that they impute a *set* of models to the decision maker.[3] How should that set be specified? Robust control theory gives a convenient way to specify and measure a set of priors starting from a single approximating model.

3 Three ordinary control problems

⟨overview⟩

This section describes three ordinary control problems. In each of them, a single decision maker maximizes an intertemporal return function by choosing a controlled stochastic process. The first two are different representations of the same underlying problem. They are cast on different probability spaces and express different timing protocols. The third, called the risk-sensitive control problem, alters the objective function of the decision maker to induce more aversion to risk.

3.1 Benchmark problem

We start with two versions of a benchmark stochastic optimal control problem. The first formulation is defined in terms of a state vector x, an underlying probability space $(\Omega, \mathcal{F}, P)$, a d-dimensional, standard Brownian motion $\{B_t : t \geq 0\}$ defined on that space, and $\{\mathcal{F}_t : t \geq 0\}$, the completion of the filtration generated by the Brownian motion B. For any stochastic process $\{a_t : t \geq 0\}$, we use a or $\{a_t\}$ to denote the process and a_t to denote the time t-component of that process. The random vector a_t maps Ω into a set $\check{A}$; $\check{a}$ denotes an element in $\check{A}$. Actions of the decision-maker form a progressively measurable stochastic process $\{c_t : t \geq 0\}$, which means that the time t component c_t is $\mathcal{F}_t$ measurable.[4] Let U be an instantaneous utility function and C be the set of admissible control processes.

⟨def1⟩ **Definition 3.1.** *The* benchmark *control problem is:*

$$J(x_0) = \sup_{c \in C} E\left[\int_0^\infty \exp(-\delta t) U(c_t, x_t) dt\right] \tag{11}$$ bench1

where the maximization is subject to

$$dx_t = \mu(c_t, x_t)dt + \sigma(c_t, x_t)dB_t \tag{12}$$ stateevolve

and where x_0 is a given initial condition.

We restrict μ and σ so that any progressively measurable control c in C implies a progressively measurable state vector process x and maintain

[3] For applied work, an attractive feature of rational expectations is that by equating the equilibrium of the model itself to the decision maker's prior, decision makers' beliefs contribute no free parameters.

[4] Progressive measurability requires that we view $c \doteq \{c_t : t \geq 0\}$ as a function of (t, ω). For any $t \geq 0$, $c : [0, t] \times \Omega$ must be measurable with respect to $\mathcal{B}_t \times \mathcal{F}_t$ where $\mathcal{B}_t$ is a collection of Borel subsets of $[0, t]$. See Karatzas and Shreve (1991) pages 4 and 5 for a discussion.

⟨a:finite⟩ **Assumption 3.2.** *$J(x_0)$ is finite.*

We shall refer to the law of motion (12) or the probability measure over sequences that it induces as the decision maker's *approximating model*. The benchmark control problem treats the approximating model as correct.

3.1.1 A static version of the benchmark problem

It is useful to restate the benchmark problem in terms of the probability space that the Brownian motion induces over continuous functions of time, thereby converting it into a static problem that pushes the state x into the background. At the same time it puts the induced probability distribution in the foreground and features the linearity of the objective in the induced probability distribution.

The d-dimensional Brownian motion B induces a multivariate Wiener measure q^0 on a *canonical space* $(\Omega^*, \mathcal{F}^*)$, where Ω^* is the space of continuous functions $f : [0, +\infty) \to \mathbb{R}^d$ and $\mathcal{F}_t^*$ is the Borel sigma algebra for the restriction of the continuous functions f to $[0, t]$. Define open sets using the sup-norm over each interval. Notice that $\iota_s(f) = f(s)$ is $\mathcal{F}_t^*$ measurable for each $0 \le s \le t$. Let $\mathcal{F}^*$ be the smallest sigma algebra containing $\mathcal{F}_t^*$ for $t \ge 0$. An event in $\mathcal{F}_t^*$ restricts continuous functions on the finite interval $[0, t]$. For any probability measure q on $(\Omega^*, \mathcal{F}^*)$, let q_t denote the restriction to $\mathcal{F}_t{}^*$. In particular, q_t^0 is the multivariate Wiener measure over the event collection $\mathcal{F}_t{}^*$.

Given a progressively measurable control c, solve the stochastic differential equation (12) to obtain a progressively measurable utility process

$$U(c_t, x_t) = \upsilon_t(c, B)$$

where $\upsilon(c, \cdot)$ is a progressively measurable family defined on $(\Omega^*, \mathcal{F}^*)$. This notation accounts for but conceals the evolution of the state vector x_t. A realization of the Brownian motion is a continuous function. Putting a probability measure q^0 on the space of continuous functions allows us to evaluate expectations. We leave implicit the dependence on B and represent the decision maker's objective as $\int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t^0 \right) dt$.

⟨staticbench⟩ **Definition 3.3.** *A* static benchmark *control problem is*

$$\tilde{J} = \max_{c \in C} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t^0 \right) dt.$$

3.1.2 Recursive version of the benchmark problem

The problem in definition 3.1 asks the decision maker once and for all at time 0 to choose an entire process $c \in C$. To transform the problem into one in which the decision maker chooses sequentially, we impose additional structure on the choice set C by restricting $\check{c}$ to be in some set $\check{C}$ that is common for all dates. With this specification of controls, we make the problem recursive by asking the decision maker to choose $\check{c}$ as a function of the state x at each date.

Definition 3.4. *The* Bellman equation *for the benchmark problem is*

$$\delta J(\check{x}) = \max_{\check{c}\in\check{C}} U(\check{c},\check{x}) + \mu(\check{c},\check{x}) \cdot J_x(\check{x}) + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})' J_{xx}(\check{x})\sigma(\check{c},\check{x})\right]. \qquad (13)$$

bell1

The recursive version of the benchmark problem (13) puts the state x_t front and center. A decision rule $c_t = \zeta_c(x_t)$ attains the right side of the Bellman equation (13).

Although the static and recursive versions of the benchmark control problem yield identical formulas for (c, x) as a function of the Brownian motion B, they differ in how they represent the same approximating model: as a stochastic differential equation in the recursive problem and as a probability distribution in the static problem. Both versions of the benchmark problem treat the decision maker's approximating model as true.

3.2 Risk-sensitive control

⟨subs:risk⟩ Let ρ be an intertemporal return function. Instead of maximizing $E\rho$ (where E continues to mean mathematical expectation), risk-sensitive control theory maximizes $E[\exp(\theta^{-1}\rho)]$, where θ^{-1} is a risk-sensitivity parameter. Risk sensitive optimal control was initiated by Jacobson (1973) and Whittle (1981) in the context of discrete-time linear-quadratic decision problems. Jacobson and Whittle showed that the risk-sensitive control law can be computed by solving a robust penalty problem of the type we have studied here.

A *risk-sensitive* control problem treats the decision maker's approximating model as true but alters preferences by appending an additional term to the right side of the Bellman equation (13):

$$\begin{aligned}\delta S(\check{x}) &= \max_{\check{c}\in\check{C}} U(\check{c},\check{x}) + \mu(\check{c},\check{x}) \cdot S_x(\check{x}) + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})' S_{xx}(\check{x})\sigma(\check{c},\check{x})\right] \\ &\quad - \frac{1}{2\theta} S_x(\check{x})'\sigma(\check{c},\check{x})\sigma(\check{c},\check{x})' S_x(\check{x}), \qquad (14)\end{aligned}$$

bell2

where $\theta > 0$. The term

$$\mu(\check{c},\check{x}) \cdot S_x(\check{x}) + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})' S_{xx}(\check{x})\sigma(\check{c},\check{x})\right]$$

in Bellman equation (14) is the local mean or dt contribution to the value function process $\{S(x_t) : t \geq 0\}$. Thus (14) coincides with the Bellman equation for the benchmark control problem (11), (12), with the additional term $-\frac{1}{2\theta}S_x(\check{x})'\sigma(\check{c},\check{x})\sigma(\check{c},\check{x})'S_x(\check{x})$ included. Notice that $S_x(x_t)'\sigma(c_t, x_t)dB_t$ gives the local Brownian contribution to the value function process $\{S(x_t) : t \geq 0\}$. The additional term in the Bellman equation is the negative of the local variance of the continuation value weighted by $\frac{1}{2\theta}$. When $\theta = +\infty$, this collapses to the benchmark control problem. When $\theta < \infty$, we call it a risk-sensitive control problem with $-\frac{1}{\theta}$ being the *risk-sensitivity* parameter. A solution of the risk-sensitive control problem is attained by a policy function

$$c_t = \alpha_c(x_t) \qquad (15)$$

cdecis1

whose dependence on θ is understood.

James (1992) studied a continuous-time, nonlinear diffusion formulation of a risk-sensitive control problem. Risk-sensitive control theory typically focuses on the case in which the discount rate δ is zero. Hansen and Sargent (1995) showed how to introduce discounting while preserving much of the mathematical structure for the linear-quadratic, Gaussian risk-sensitive control problem. They applied the recursive utility framework developed by Epstein and Zin (1989) in which the risk-sensitive adjustment is applied recursively to the continuation values. Recursive formulation (14) gives the continuous-time counterpart for Markov diffusion processes. Duffie and Epstein (1992) characterized the preferences that underlie this specification.

4 Fear of model misspecification

⟨fear⟩ For a given θ, the optimal risk-sensitive decision rule emerges from other problems in which the decision maker's objective function remains the one described in the benchmark problem (11) and in which the adjustment to the continuation value in (14) reflects not altered preferences but distrust of the model (12). Moreover, just as we could have formulated the benchmark problem either as a static problem with induced distributions or as a recursive problem, there are also static and recursive representations of robust control problems.

Each of our decision problems for promoting robustness to model misspecification is a zero-sum, two-player game in which a maximizing player ('the decision maker') chooses a best response to a malevolent player ('nature') who can alter the stochastic process within prescribed limits. The minimizing player's malevolence is the maximizing player's tool for analyzing the fragility of alternative decision rules. Each game uses a Nash equilibrium concept. We portray games that differ from one another in three dimensions: (1) the protocols that govern the timing of players' decisions, (2) the constraints on the malevolent player's choice of models; and (3) the mathematical spaces in terms of which the games are posed. Because the state spaces and probability spaces on which they are defined differ, the recursive versions of these problems yield decision rules that differ from (15). Despite that, all of the formulations give rise to identical decision processes for c, all of which in turn are equal to those that apply the optimal risk sensitive decision rule (15) to the transition equation (12).

The equivalence of their outcomes provides interesting alternative perspectives from which to understand the decision maker's response to possible model misspecification.[5] That outcomes are identical for these different games means that when all is said and done, the timing protocols don't matter. Because some of the timing protocols correspond to 'static' games while others enable sequential choices, equivalence of equilibrium outcomes implies a form of dynamic consistency.

Jacobson (1973) and Whittle (1981) first showed that the risk-sensitive control law can be computed by solving a robust penalty problem of the type we have studied here, but without

[5] See section 9 of Anderson, Hansen, and Sargent (2003) for an application.

discounting. Subsequent research reconfirmed this link in nonsequential and undiscounted problems where the robust problem was typically posed in a nonstochastic environment. Petersen, James, and Dupuis (2000) explicitly consider an environment with randomness, but do not make the link to recursive risk-sensitivity.

5 Two robust control problems defined on sets of probability measures

⟨secdatezero⟩ We formalize the connection between two problems that are robust counterparts to the static version of the benchmark control problem (3.3). In contrast to the benchmark problem, we do not fix an induced probability distribution q^o and optimize. Instead these games express alternative *models* as alternative induced probability distributions. One of the two games falls naturally into the framework of Gilboa and Schmeidler (1989) and the other is closely linked to risk-sensitive control. An advantage of working with the induced distributions is that a requisite convexity is easy to establish and exploit in establishing the connection between the two games.

5.1 Entropy and absolute continuity over finite intervals

In formulating these problems, we settle on a particular notion of absolute continuity of one infinite-time stochastic process with respect to another. We use something weaker than the standard notion that characterizes two stochastic processes as being absolutely continuous with respect to each other if they agree about 'tail events'. Roughly speaking, the weaker concept requires that the two measures being compared both put positive probability on all of the same events, *except* tail events. This weaker notion of absolute continuity is interesting for applied work because of what it implies about how quickly it is possible statistically to distinguish one model from another.

Recall that the Brownian motion B induces a multivariate Wiener measure on $(\Omega^*, \mathcal{F}^*)$ that we have denoted q^0. For any probability measure q on $(\Omega^*, \mathcal{F}^*)$, we let q_t denote the restriction to ${\mathcal{F}_t}^*$. In particular, q_t^0 is the multivariate Wiener measure over the event ${\mathcal{F}_t}^*$.

Definition 5.1. *A distribution q is said to be* absolutely continuous over finite intervals *with respect to q^0 if q_t is absolutely continuous with respect to q_t^0 for all t.*[6]

Let Q be the set of all distributions that are absolutely continuous with respect to q^0 over finite intervals. The set Q is convex. Absolute continuity over finite intervals captures the idea that two models are difficult to distinguish from samples of finite length. If q is absolutely continuous with respect to q^0 over finite intervals, we can construct likelihood

[6] Kabanov, Lipcer, and Sirjaev (1979) refer to this concept as *local* absolute continuity. Although Kabanov, Lipcer, and Sirjaev (1979) define local absolute continuity through the use of stopping times, they argue that their definition is equivalent to this "simpler one".

ratios for finite histories at any calendar date t. To measure the discrepancy between models over an infinite horizon, we use a discounted measure of *relative entropy*:

$$\tilde{\mathcal{R}}(q) \doteq \delta \int_0^\infty \exp(-\delta t) \left(\int \log \left(\frac{dq_t}{dq_t^0} \right) dq_t \right) dt, \tag{16}$$

RQ

where $\frac{dq_t}{dq_t^0}$ is the Radon-Nikodym derivative of q_t with respect to q_t^0. This discrepancy measure is convex in q as shown in the appendix B. (See claim B.1.)

The distribution q is absolutely continuous with respect to q^0 when

$$\int \log \left(\frac{dq}{dq^0} \right) dq < +\infty.$$

In this case a law of large numbers that applies under q_0 must also apply under q, so that discrepancies between them are at most 'temporary'. We introduce discounting in part to provide an alternative interpretation of the recursive formulation of risk-sensitive control as expressing a fear of model misspecification rather than extra aversion to well understood risks. By discounting in (16) we allow

$$\int \log \left(\frac{dq}{dq^0} \right) dq = +\infty. \tag{17}$$

req1

Time series averages of functions that converge almost surely under q^0 can converge to a different limit under q, or they may not converge at all. That would allow a statistician to distinguish q from q^0 with a continuous record of data on an infinite interval.[7] But we want these alternative models to be close enough to the approximating model that they are statistically difficult to distinguish from it after having observed a continuous data record of only finite length N on the state. We implement this requirement by insisting that (16) holds.

The presence of discounting in (16) and its absence from (17) are significant. By directing the decision maker's attention to alternative models that satisfy (17), we make the decision maker seek robustness against models that can be distinguished from the approximating model with an infinite data record, but because they also satisfy (16), it is difficult to distinguish them from a finite data record. We have in mind settings of δ for which impatience outweighs the decision maker's ability eventually to learn specifications that give superior fits, prompting him to focus on designing a robust decision rule.

We now have the vocabulary to state two static robust control problems that use Q as a family of distortions to the probability distribution q^0 in the benchmark problem:

ef:multiplier⟩

[7] Our specification allows Q measures to put different probabilities on tail events, which prevents the conditional measures from merging as Blackwell and Dubins (1962) show will occur under absolute continuity. See Kalai and Lerner (1993) and Jackson, Kalai, and Smordoninsky (1999) for implications of absolute continuity for learning.

Definition 5.2. *A nonsequential penalty robust control problem is*

$$\tilde{V}(\theta) = \max_{c \in C} \min_{q \in Q} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt + \theta \tilde{\mathcal{R}}(q).$$

ef:constraint⟩ **Definition 5.3.** *A nonsequential constraint robust control problem is*

$$\tilde{K}(\eta) = \max_{c \in C} \min_{q \in Q(\eta)} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt$$

where $Q(\eta) = \{q \in Q : \tilde{\mathcal{R}}(q) \leq \eta\}$.

The first problem is closely linked to the risk sensitive control problem. The second problem fits into the max-min expected utility or multiple priors model advocated by Gilboa and Schmeidler (1989), the set of priors being $Q(\eta)$. We use a set of θ's to index a family of penalty robust control problems and a set of η's to index a family of constraint robust control problems. The two types of problems are linked by the Lagrange multiplier theorem.

5.2 Relation between the constraint and penalty problems

This subsection establishes two important things about the two static multiple priors problems 5.2 and 5.3: (1) we show that we can interpret the robustness parameter θ in problem 5.2 as a Lagrange multiplier on the specification-error constraint $\tilde{\mathcal{R}}(q) \leq \eta$ in problem 5.3;[8] (2) we display technical conditions that make the solutions of the two problems equivalent to one another. We shall exploit both of these results in later sections.

The simultaneous maximization and minimization means that the link between the penalty and constraint problem is not a direct implication of the Lagrange multiplier Theorem. The following treatment exploits convexity of $\mathcal{R}^*$ in Q. The analysis follows that in Petersen, James, and Dupuis (2000), although our measure of entropy differs.[9] As in Petersen, James, and Dupuis (2000), we use tools of convex analysis contained in Luenberger (1969) to establish the connection between the two problems.

Assumption 3.2 makes optimized objectives for both the penalty and constraint robust control problems be less than $+\infty$. They can be $-\infty$, depending on the magnitudes of θ and η.

Given an $\eta^* > 0$, add $-\theta\eta^*$ to the objective in problem 5.2, which for given θ has no impact on the control law.[10] For a given c, the objective of the constraint robust control problem is linear in q and the entropy measure $\tilde{\mathcal{R}}$ in the constraint is convex in q. Moreover,

[8] This connection is regarded as self-evident throughout the literature on robust control. It has been explored in the context of a linear-quadratic control problem, informally by Hansen, Sargent, and Tallarini (1999), and formally by Hansen and Sargent (2005).

[9] To accommodate discounting in the recursive, risk sensitive control problem, we are led to include discounting in our measure of entropy. See appendix B.

[10] However, it will alter which θ results in the highest objective.

the family of admissible probability distributions Q is itself convex. Thus we formulate the constraint version of the robust control problem (problem 5.3) as a Lagrangian:

$$\max_{c\in C}\min_{q\in Q}\max_{\theta\geq 0}\int_0^\infty \exp(-\delta t)\left(\int \upsilon_t(c)dq_t\right)dt+\theta\left[\tilde{\mathcal{R}}(q)-\eta\right].$$

The optimizing multiplier θ is degenerate for many choices of q: it is infinite if q violates the constraint and zero if the constraint is slack. Therefore, we include $\theta=+\infty$ in the choice set for θ. Exchanging the order of $\max_\theta$ and $\min_q$ attains the same value of q. The Lagrange multiplier theorem allows us to study:

$$\max_{c\in C}\max_{\theta\geq 0}\min_{q\in Q}\int_0^\infty \exp(-\delta t)\left(\int \upsilon_t(c)dq_t\right)dt+\theta\left[\tilde{\mathcal{R}}(q)-\eta\right]. \tag{18}$$ LMproblem

A complication in the argument arises at this point because the maximizing θ in (18) depends on the choice of c. In solving a robust control problem, we are most interested in the c that solves the constraint robust control problem. We can find the appropriate choice of θ by changing the order of $\max_c$ and $\max_\theta$ to obtain:

$$\max_{\theta\geq 0}\max_{c\in C}\min_{q\in Q}\int_0^\infty \exp(-\delta t)\left(\int \upsilon_t(c)dq_t\right)dt+\theta\left[\tilde{\mathcal{R}}(q)-\eta^*\right]=\max_{\theta\geq 0}\tilde{V}(\theta)-\theta\eta^*$$

since for a given θ the term $-\theta\eta^*$ does not effect the extremizing choices of (c,q).

⟨claim1⟩ **Claim 5.4.** *For $\eta^*>0$, suppose that c^* and q^* solve the constraint robust control problem. Then there exists a $\theta^*>0$ such that the corresponding penalty robust control problem has the same solution. Moreover,*

$$\tilde{K}(\eta^*)=\max_{\theta\geq 0}\tilde{V}(\theta)-\theta\eta^*.$$

Proof. This result is essentially the same as Theorem 2.1 of Petersen, James, and Dupuis (2000) and follows directly from Luenberger (1969). □

This claim gives $\tilde{K}$ as the Legendre transform of $\tilde{V}$. Moreover, by adapting an argument of (Luenberger 1969), we can show that $\tilde{K}$ is decreasing and convex in η.[11] In what follows, we shall be interested in recovering $\tilde{V}$ from $\tilde{K}$ as the inverse Legendre transform via:

$$\tilde{V}(\theta^*)=\min_{\eta\geq 0}\tilde{K}(\eta)+\theta^*\eta. \tag{19}$$ inverse

It remains to justify this recovery formula.

We call *admissible* those nonnegative values of θ for which it is feasible to make the objective function greater than $-\infty$. If $\hat{\theta}$ is admissible, values of θ larger than $\hat{\theta}$ are also

[11]This follows because we may view $\tilde{K}$ as the maximum over convex functions indexed by alternative consumption processes.

admissible, since these values only make the objective larger. Let $\underline{\theta}$ denote the greatest lower bound for admissible values of θ. Consider a value $\theta^* > \underline{\theta}$. Our aim is to find a constraint associated with this choice of θ.

It follows from claim 5.4 that

$$\tilde{V}(\theta^*) \leq \tilde{K}(\eta) + \theta^* \eta$$

for any $\eta > 0$ and hence

$$\tilde{V}(\theta^*) \leq \min_{\eta \geq 0} \tilde{K}(\eta) + \theta^* \eta.$$

Moreover,

$$\tilde{K}(\eta) \leq \min_{q \in Q(\eta)} \max_{c \in C} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt,$$

since making the maximization the inner problem instead of minimization can only increase the objective. Thus

$$\begin{aligned} \tilde{V}(\theta^*) &\leq \min_{\eta \geq 0} \left[\min_{q \in Q(\eta)} \max_{c \in C} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt + \theta^* \eta \right] \\ &= \min_{\eta \geq 0} \left[\min_{q \in Q(\eta)} \max_{c \in C} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt + \theta^* \tilde{R}(q) \right] \\ &= \min_{q \in Q} \max_{c \in C} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt + \theta^* \tilde{R}(q). \end{aligned}$$

For the second equality, the minimization over η is important. We avoid showing that the worst-case distribution in the set $Q(\eta)$ is a q for which $\tilde{\mathcal{R}}(q) = \eta$. If instead the worst case q satisfies $\tilde{\mathcal{R}}(q) < \tilde{\eta}$ for some $\tilde{\eta}$, we know that equality must hold for a smaller value of η. Notice that the last equality gives a $\min - \max$ analogue to the *nonsequential penalty problem* (5.2) except with the role of minimization and maximization reversed. If the resulting value continues to be $\tilde{V}(\theta^*)$, we have verified (19).

We shall invoke the following assumption:

Assumption 5.5. *For $\theta > \underline{\theta}$*

⟨exchange⟩

$$\begin{aligned} \tilde{V}(\theta) &= \max_{c \in C} \min_{q \in Q} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt + \theta \tilde{\mathcal{R}}(q) \\ &= \min_{q \in Q} \max_{c \in C} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t \right) dt + \theta \tilde{\mathcal{R}}(q). \end{aligned}$$

Both equalities assume that the maximum and minimum are attained. Because minimization occurs first, without the assumption the second equality has to be replaced by a less than or equal sign ($\leq$).

⟨claim2⟩ **Claim 5.6.** *Suppose that Assumption 5.5 is satisfied and that for $\theta^* > \underline{\theta}$, c^* is the maximizing choice of c for the penalty robust control problem 5.2. Then that c^* also solves the constraint robust control problem 5.3 for $\eta^* = \tilde{\mathcal{R}}(q^*)$ where η^* solves*

$$\tilde{V}(\theta^*) = \min_{\eta \geq 0} \tilde{K}(\eta) + \theta^* \eta.$$

Since $\tilde{K}$ is decreasing and convex, $\tilde{V}$ is increasing and concave in θ. The Legendre and inverse Legendre transforms given in claims 5.4 and 5.6 fully describe the mapping between the constraint index η^* and the penalty parameter θ^*. However, given η^*, they do not imply that the associated θ^* is unique, nor for a given $\theta^* > \underline{\theta}$ do they imply that the associated η^* is unique.

While claim 5.6 maintains assumption 5.5, claim 5.4 does not. Without assumption 5.5, we do not have a proof that $\tilde{V}$ is concave. Moreover, for some values of θ^* and a solution pair (c^*, q^*) of the penalty problem, we may not be able to produce a corresponding constraint problem. Nevertheless, the family of penalty problems (indexed by θ) continues to embed the solutions to the constraint problems (indexed by η) as justified by claim 5.4. We are primarily interested in problems for which assumption 5.5 is satisfied and will subsequently provide some sufficient conditions for this assumption. One reason for interest in this assumption is given in the next subsection.

5.3 Bayesian interpretation of outcome of static game

?⟨bayes22⟩?

A widespread device for interpreting a statistical decision rule is to find a probability distribution for which the rule is optimal. Here we seek an induced probability distribution for B such that the solution for c from either the constraint or penalty robust decision problem is optimal for a counterpart to the benchmark problem. When we can produce such a distribution, we say that we have a Bayesian interpretation for the robust decision rule. (See Blackwell and Girshick (1954) and Chamberlain (2000) for related discussions.)

The freedom to exchange orders of maximization and minimization in problem 5.3 justifies such a Bayesian interpretation of the decision process $c \in C$ that solves game 5.2. Let (c^*, q^*) be the equilibrium of game 5.2. Given the worst case model q^*, consider the control problem:

$$\max_{c \in C} \int_0^\infty \exp(-\delta t) \left(\int \upsilon_t(c) dq_t^* \right) dt. \tag{20}$$

bayes50

Problem (20) is a version of our static benchmark problem 3.3 with a fixed model q^* that is distorted relative to the approximating model q^0. The optimal choice of a progressively measurable c takes q^* as exogenous. The optimal decision c^* is not altered by adding $\theta\tilde{\mathcal{R}}(q^*)$ to the objective. Therefore, being able to exchange orders of extremization in 5.2 allows us to support a solution to the penalty problem by a particular distortion in the Weiner measure. The implied least favorable q^* assigns a different (induced) probability measure for the exogenous stochastic process $\{B_t : t \geq 0\}$. Given that distribution, c^* is the ordinary (non robust) optimal control process.

Having connected the penalty and the constraint problem, in what follows we will focus primarily on the penalty problem. For notational simplicity, we will simply fix a value of θ and not formally index a family of problems by this parameter value.

6 Games on fixed probability spaces

⟨secperturb⟩ This section focuses on important technical details that are involved in moving from the static to the recursive versions of the multiple probability games 5.2 and 5.3. In so doing it is convenient to represent alternative model specifications as martingale 'preference shocks' on a common probability space. This will allow us to formulate the games as two-player zero-sum differential games and to appeal to some existing results for such games. Instead of working with multiple distributions on the measurable space $(\Omega^*, \mathcal{F}^*)$, we now use the original probability space $(\Omega, \mathcal{F}, P)$ in conjunction with nonnegative martingales. This section presents a convenient way to parameterize the martingales and issues a caveat about this parameterization.

6.1 Martingales and finite interval absolute continuity

For any continuous function f in Ω^*, let

$$\begin{aligned} \kappa_t(f) &= \left(\frac{dq_t}{dq_t^0}\right)(f) \\ z_t &= \kappa_t(B) \end{aligned} \qquad (21)$$

radon

where κ_t is the Radon-Nikodym derivative of q_t with respect to q_t^0.

⟨qzequiv⟩ **Claim 6.1.** *Suppose that for all $t \geq 0$, q_t is absolutely continuous with respect to q_t^0. The process $\{z_t : t \geq 0\}$ defined via (21) on $(\Omega, \mathcal{F}, P)$ is a nonnegative martingale adapted to the filtration $\{\mathcal{F}_t : t \geq 0\}$ with $Ez_t = 1$. Moreover,*

$$\int \phi_t dq_t = E\left[z_t \phi_t(B)\right] \qquad (22)$$

mart

for any bounded and $\mathcal{F}_t^*$ measurable function ϕ_t. Conversely, if $\{z_t : t \geq 0\}$ is a nonnegative progressively measurable martingale with $Ez_t = 1$, then the probability measure q defined via (22) is absolutely continuous with respect to q^0 over finite intervals.

Proof. The first part of this claim follows directly from the proof of theorem 7.5 in Liptser and Shiryaev (2000). Their proof is essentially a direct application of the Law of Iterated Expectations and the fact that probability distributions necessarily integrate to one. Conversely, suppose that z is a nonnegative martingale on $(\Omega, \mathcal{F}, P)$ with unit expectation. Let ϕ_t be any nonnegative, bounded and $\mathcal{F}_t^*$ measurable function. Then (22) defines a measure because indicator functions are nonnegative, bounded functions. Clearly $\int \phi_t dq_t = 0$ whenever $E\phi_t(B) = 0$. Thus q_t is absolutely continuous with respect to q_t^0, the measure induced by Brownian motion restricted to $[0, t]$. Setting $\phi_t = 1$ shows that q_t is in fact a probability measure for any t. □

Claim 6.1 is important because it allows us to integrate over $(\Omega^*, \mathcal{F}^*, q)$ by integrating against a martingale z on the original probability space $(\Omega, \mathcal{F}, P)$.

6.2 Representing martingales

By exploiting the Brownian motion information structure, we can attain a convenient representation of a martingale. Any martingale z with a unit expectation can be portrayed as

$$z_t = 1 + \int_0^t k_u dB_u$$

where k is a progressively measurable d-dimensional process that satisfies:

$$P\left\{\int_0^t |k_u|^2 du < \infty\right\} = 1$$

for any finite t (see Revuz and Yor (1994), Theorem V.3.4). Define:

$$h_t = \begin{cases} k_t/z_t & \text{if} \quad z_t > 0 \\ 0 & \text{if} \quad z_t = 0. \end{cases} \tag{23}$$ hconstruct

Then z solves the integral equation

$$z_t = 1 + \int_0^t z_u h_u dB_u$$

and its differential counterpart

$$dz_t = z_t h_t dB_t \tag{24}$$ martdiff

with initial condition $z_0 = 1$, where for $t > 0$

$$P\left\{\int_0^t (z_u)^2 |h_u|^2 du < \infty\right\} = 1.$$

The scaling by $(z_u)^2$ permits

$$\int_0^t |h_u|^2 du = \infty$$

provided that $z_t = 0$ on this event.

In reformulating the nonsequential penalty problem 5.2, we parameterize nonnegative martingales by progressively measurable processes h. We introduce a new state z_t initialized at one, and take h to be under the control of the minimizing agent.

6.3 Representing likelihood ratios

We are now equipped to fill in some important details associated with using martingales to represent likelihood ratios for dynamic models. Before addressing these issues, we use a simple static example to exhibit an important idea.

6.3.1 A static example

The static example is designed to illustrate two alternative ways of representing the expected value of a likelihood ratio by changing the measure with respect to which it is evaluated. Consider two models of a vector y. In the first, y is normally distributed with mean ν and covariance matrix I. In the second, y is normally distributed with mean zero and covariance matrix I. The logarithm of the ratio of the first density to the second is:

$$\ell(y) = \left(\nu \cdot y - \frac{1}{2}\nu \cdot \nu\right).$$

Let E^1 denote the expectation under model one and E^2 under model two. Properties of the log-normal distribution imply that

$$E^1 \exp\left[\ell(y)\right] = 1.$$

Under the second model

$$E^2\ell(y) = E^1\ell(y)\exp[\ell(y)] = \frac{1}{2}\nu \cdot \nu,$$

which is relative entropy.

6.3.2 The dynamic counterpart

We now consider the dynamic counterpart to the static example by showing two ways to represent likelihood ratios, one under the original Brownian motion model and another under the model associated with a nonnegative martingale z. First we consider the likelihood ratio under the Brownian motion model for B. As noted above, the solution to (24) can be represented as an exponential:

$$z_t = \exp\left(\int_0^t h_u \cdot dB_u - \frac{1}{2}\int_0^t |h_u|^2 du\right). \tag{25}$$ likelihood

We allow $\int_0^t |h_u|^2 du$ to be infinite with positive probability and adopt the convention that the exponential is zero when this event happens. In the event that $\int_0^t |h_u|^2 du < \infty$, we can define the stochastic integral $\int_0^t h_u dB_u$ as an appropriate probability limit (see Lemma 6.2 of Liptser and Shiryaev (2000)).

When z is a martingale, we can interpret the right side of (25) as a formula for the likelihood ratio of two models evaluated under the Brownian motion specification for B. Taking logarithms, we find that

$$\ell_t = \int_0^t h_u \cdot dB_u - \frac{1}{2}\int_0^t |h_u|^2 du.$$

Since h is progressively measurable, we can write:

$$h_t = \psi_t(B).$$

Changing the distribution of B in accordance with q gives another characterization of the likelihood ratio. The Girsanov Theorem implies

claimperturb⟩? **Claim 6.2.** *If for all $t \geq 0$, q_t is absolutely continuous with respect to q_t^0, then q is the induced distribution for a (possibly weak) solution B to a stochastic differential equation defined on a probability space $(\Omega, \mathcal{F}, \tilde{P})$:*

$$dB_t = \psi_t(B)dt + d\tilde{B}_t \tag{26}$$?weak?

for some progressively measurable ψ defined on $(\Omega^*, \mathcal{F}^*)$ and some Brownian motion $\tilde{B}$ that is adapted to $\{\mathcal{F}_t : t \geq 0\}$. Moreover, for each t

$$\tilde{P}\left[\int_0^t |\psi_u(B)|^2 du < \infty\right] = 1.$$

Proof. ¿From Lemma 6.1 there is a nonnegative martingale z associated with the Radon-Nikodym derivative of q_t with respect to q_t^0. This martingale has expectation unity for all t. The conclusion follows from a generalization of the Girsanov Theorem (*e.g.* see Liptser and Shiryaev (2000) Theorem 6.2). □

The $\psi_t(B)$ is the same as that used to represent h_t where h is defined by (23). Under the distribution $\tilde{P}$,

$$B_t = \int_0^t h_u du + \tilde{B}_t$$

where $\tilde{B}_t$ is a Brownian motion with respect to the filtration $\{\mathcal{F}_t : t \geq 0\}$. In other words, we obtain perturbed models by replacing the Brownian motion model for a shock process with a Brownian motion with a drift.

Using this representation, we can write the logarithm of the likelihood ratio as:

$$\tilde{\ell}_t = \int_0^t \psi_u(B) \cdot d\tilde{B}_u + \frac{1}{2}\int_0^t |\psi_u(B)|^2 du. \tag{27}$$?likelihood2?

Claim 6.3. *For $q \in Q$, let z be the nonnegative martingale associated with q and let h be the progressively measurable process satisfying (49). Then*

$$\tilde{\mathcal{R}}(q) = \frac{1}{2}E\left[\int_0^\infty \exp(-\delta t) z_t |h_t|^2 dt\right].$$

Proof. See appendix B. □

This claim leads us to define a discounted entropy measure for nonnegative martingales:

$$\mathcal{R}^*(z) \doteq \frac{1}{2}E\left[\int_0^\infty \exp(-\delta t) z_t |h_t|^2 dt\right]. \tag{28}$$ entconst

6.4 A martingale version of a robust control problem

Modelling alternative probability distributions by introducing preference shocks that are martingales on a common probability space is mathematically convenient. This approach leads us to reformulate the penalty robust control problem (problem 5.2) as:

f:multiplier2⟩ **Definition 6.4.** *A nonsequential martingale robust control problem is*

$$\max_{c \in C} \min_{h \in H} E \left(\int_0^\infty \exp(-\delta t) z_t \left[U(c_t, x_t) + \frac{\theta}{2} |h_t|^2 \right] d\,t \right) \tag{29}$$ zobjective

subject to:

$$\begin{aligned} dx_t &= \mu(c_t, x_t) + \sigma(c_t, x_t) dB_t \\ dz_t &= z_t h_t \cdot dB_t. \end{aligned} \tag{30}$$ zstateevolve

But there is potentially a technical problem with this formulation. There may exist control process h and corresponding processes z such that z is a nonnegative local martingale for which $\mathcal{R}^*(z) < \infty$ yet z is not a martingale. We have not ruled out nonnegative supermartingales that happen to be local martingales. This means that even though z is a local martingale, it might satisfy only the inequality

$$E\left(z_t | \mathcal{F}_s\right) \leq z_s$$

for $0 < s \leq t$. Even when we initialize z_0 to one, z_t may have a mean less than one and the corresponding measure will not be a probability measure. Then we would have given the minimizing agent more options than we intend.

For this not to cause difficulty, at the very least we have to show that the minimizing player's choice of h in problem 6.4 is associated with a z that is a martingale and not just a supermartingale.[12] More generally, we have to verify that enlarging the set of processes z as we have done does not alter the equilibrium of the two-player zero-sum game. In particular, consider the second problem in assumption 5.5. It suffices to show that the minimizing h implies a z that is a martingale. If we assume that condition 5.5 is satisfied, then it suffices to check this for the following timing protocol:

$$\min_{h \in H} \max_{c \in C} E \left(\int_0^\infty \exp(-\delta t) z_t \left[U(c_t, x_t) + \frac{\theta}{2} |h_t|^2 \right] d\,t \right)$$

subject to (30) with initial conditions $x_0, z_0 = 1$.[13] In appendix C, we show how to establish that the solution is indeed a martingale.

[12] Alternatively, we might interpret the supermartingale as allowing for an escape to a terminal absorbing state with a terminal value function equal to zero. The expectation of z_t gives the probability that an escape has not happened as of date t. The existence of such terminal state is not, however, entertained in our formulation of 5.2.

[13] To see this let $H^* \subseteq H$ be the set of controls h for which z is a martingale and let $\text{obj}(h, c)$ be the

7 Sequential timing protocol for a penalty formulation

⟨Rmult⟩

The martingale problem 6.4 assumes that at time zero both decision makers commit to decision processes whose time t components are measurable functions of $\mathcal{F}_t$. The minimizing decision maker who chooses distorted beliefs h takes c as given; and the maximizing decision maker who chooses c takes h as given. Assumption 5.5 asserts that the order in which the two decision makers choose does not matter,

This section studies a two-player zero-sum game with a timing protocol which both players choose sequentially. We set forth conditions that imply that with sequential choices we obtain the same time zero value function and the same outcome path that we would with a timing protocol in which both players choose once and for all at time 0. The sequential formulation is convenient computationally and also gives a way to justify the exchange of orders of extremization stipulated by assumption 5.5.

We have used c to denote the control process and $\check{c} \in \check{C}$ to denote the value of a control at a particular date. We let $\check{h} \in \check{H}$ denote the realized martingale control at any particular date. We can think of $\check{h}$ as a vector in $\mathbb{R}^d$. Similarly, we think of $\check{x}$ and $\check{z}$ as being realized states.

To analyze outcomes under a sequential timing protocol, we think of varying the initial state and define a value function $M(x_0, z_0)$ as the optimized objective function (29) for the martingale problem. By appealing to results of Fleming and Souganidis (1989), we can verify that $\tilde{V}(\theta) = M(\check{x}, \check{z}) = \check{z}V(\check{x})$, provided that x is initialized at x_0 and z is initialized at one. Under a sequential timing protocol, this same value function gives the *continuation value* for evaluating states reached at subsequent time periods.

Fleming and Souganidis (1989) show that a *Bellman-Isaacs condition* renders equilibrium outcomes under two-sided commitment at date zero identical with outcomes of a Markov perfect equilibrium in which the decision rules of both agents are chosen sequentially, each as a function of the state vector x_t. The Hamilton-Jacobi-Bellman equation for the infinite-horizon zero-sum two-player martingale game is:

$$\begin{aligned} \delta \check{z} V(\check{x}) \quad = \quad & \max_{\check{c} \in \check{C}} \min_{\check{h}} \check{z} U(\check{c}, \check{x}) + \check{z} \frac{\theta}{2} \check{h} \cdot \check{h} + \mu(\check{c}, \check{x}) \cdot V_x(\check{x}) \check{z} \\ & + \check{z} \frac{1}{2} \text{trace} \left[\sigma(\check{c}, \check{x})' V_{xx}(\check{x}) \sigma(\check{c}, \check{x}) \right] + \check{z} \check{h} \cdot \sigma(\check{c}, \check{x})' V_x(\check{x}) \end{aligned} \tag{32}$$ HJB

objective as a function of the controls. Then under Assumption 5.5 we have

$$\min_{h \in H^*} \max_{c \in C} \text{obj}(h, c) \geq \min_{h \in H} \max_{c \imath C} \text{obj}(h, c) = \max_{c \in C} \min_{h \in H} \text{obj}(h, c) \leq \max_{c \in C} \min_{h \in H^*} \text{obj}(h, c). \tag{31}$$ toobig

If we demonstrate, the first inequality $\geq$ in (31) is an equality, it follows that

$$\min_{h \in H^*} \max_{c \in C} \text{obj}(h, c) \leq \max_{c \in C} \min_{h \in H^*} \text{obj}(h, c).$$

Since the reverse inequality is always satisfied provided that the extrema are attained, this inequality can be replaced by an equality. It follows that the second inequality $\leq$ in (31) must in fact be an equality as well.

where V_x is the vector of partial derivatives of V with respect to $\check{x}$ and V_{xx} is the matrix of second derivatives. The diffusion specification makes this Bellman equation a partial differential equation that has multiple solutions that correspond to different boundary conditions. To find the true value function and to justify the associated control laws requires that we apply a *Verification Theorem* (*e.g.* see Theorem 5.1 of Fleming and Soner (1993)).

The scaling of partial differential equation (32) by $\check{z}$ verifies our guess that the value function is linear in z. This allows us to study the alternative Hamilton-Jacobi-Bellman equation:

$$\begin{aligned}\delta V(\check{x}) &= \max_{\check{c}\in\check{C}} \min_{\check{h}} U(\check{c},\check{x}) + \frac{\theta}{2}\check{h}\cdot\check{h} + \left[\mu(\check{c},\check{x}) + \sigma(\check{c},\check{x})\check{h}\right]\cdot V_x(\check{x}) \\ &\quad + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})' V_{xx}(\check{x})\sigma(\check{c},\check{x})\right], \end{aligned} \tag{33}$$

HJB2

which involves only the $\check{x}$ component of the state vector and not $\check{z}$.[14]

A Bellman-Isaacs condition renders the order of action taken in the recursive game to be inconsequential. Formally it requires:

bellmanisaacs⟩ **Assumption 7.1.** *The value function V satisfies*

$$\begin{aligned}\delta V(\check{x}) &= \max_{\check{c}\in\check{C}} \min_{\check{h}} U(\check{c},\check{x}) + \frac{\theta}{2}\check{h}\cdot\check{h} + \left[\mu(\check{c},\check{x}) + \sigma(\check{c},\check{x})\check{h}\right]\cdot V_x(\check{x}) \\ &\quad + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})' V_{xx}(\check{x})\sigma(\check{c},\check{x})\right] \\ &= \min_{\check{h}} \max_{\check{c}\in\check{C}} U(\check{c},\check{x}) + \frac{\theta}{2}\check{h}\cdot\check{h} + \left[\mu(\check{c},\check{x}) + \sigma(\check{c},\check{x})\check{h}\right]\cdot V_x(\check{x}) \\ &\quad + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})' V_{xx}(\check{x})\sigma(\check{c},\check{x})\right] \end{aligned}$$

Appendix D describes three ways to verify this Bellman-Isaacs condition. The infinite-horizon counterpart to the result of Fleming and Souganidis (1989) asserts that the Bellman-Isaacs condition implies assumption 5.5 and hence $\tilde{V}(\theta) = V(x_0)$ because z is intialized at unity.

[14] We can construct another differential game for which V is the value function replacing dB_t by $h_t dt + dB_t$ in the evolution equation instead of introducing a martingale. In this way we would perturb the process rather than the probability distribution. While this approach can be motivated using Girsanov's Theorem, some subtle differences between the resulting perturbation game and the martingale game arise because the history of $\hat{B}_t = \int_0^t h_u du + B_t$ can generateeither a smaller or a larger filtration than that of the Brownian motion B. When it generates a smaller sigma algebra, we would be compelled to solve a combined control and filtering problem if we think of $\hat{B}$ as the generating the information available to the decision maker. If $\hat{B}$ generates a larger information set, then we are compelled to consider weak solutions to the stochastic differential equations that underlie the decision problem. Instead of extensively developing this alternative interpretation of V (as we did in an earlier draft), we simply think of the partial differential equation (33) as a means of simplifying the solution to the martingale problem.

7.1 A representation of z^*

One way to represent the worst-case martingale z^* in the recursive penalty game opens a natural transition to the risk-sensitive ordinary control problem whose Bellman equation is (14). The minimizing player's decision rule is $\check{h} = \alpha_h(\check{x})$ where

$$\alpha_h(\check{x}) = -\frac{1}{\theta}\sigma^*(\check{x})'V_x(\check{x}) \tag{34}$$ hformula

and $\sigma^*(\check{x}) \equiv \sigma^*(\alpha_c(\check{x}), \check{x})$. Suppose that $V(\check{x})$ is twice continuously differentiable. Applying the formula on page 226 of Revuz and Yor (1994), form the positive function:

$$C(\check{x}) = \exp\left[-\frac{1}{\theta}V(\check{x})\right].$$

Then

$$z_t^* = \frac{C(x_t)}{C(x_0)}\exp\left[-\int_0^t w(x_u)du\right]$$

where w is constructed to ensure that z has a zero drift. The worst case distribution assigns more weight to bad states as measured by an exponential adjustment to the value function. This representation leads directly to the risk-sensitive control problem that we take up in the next subsection.

7.2 Risk sensitivity revisited

The Bellman equation for the recursive, risk-sensitive control problem is obtained by substituting the solution (34) for h into the partial differential equation (33):

$$\begin{aligned} \delta V(\check{x}) &= \max_{\check{c}\in\check{C}} \min_{\check{h}} U(\check{c},\check{x}) + \frac{\theta}{2}\check{h}\cdot\check{h} + \left[\mu(\check{c},\check{x}) + \sigma(\check{c},\check{x})\check{h}\right]\cdot V_x(\check{x}) \\ &\quad + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})'V_{xx}(\check{x})\sigma(\check{c},\check{x})\right] \\ &= \max_{\check{c}\in\check{C}} U(\check{c},\check{x}) + \mu(\check{c},\check{x})\cdot V_x(\check{x}) \\ &\quad + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})'V_{xx}(\check{x})\sigma(\check{c},\check{x})\right] \\ &\quad - \frac{1}{2\theta}V_x(\check{x})'\sigma(\check{c},\check{x})\sigma(\check{c},\check{x})'V_x(\check{x}) \end{aligned} \tag{35}$$ hjbrisk

The value function$V$ function for the robust penalty problem is also the value function for the risk sensitive control problem of section 3.2. The *risk sensitive* interpretation excludes worries about misspecified dynamics and instead enhances the control objective with aversion to risk in a way captured by the local variance of the continuation value. While mathematically related to the situation discussed in James (1992) (see pages 403 and 404), the presence of discounting in our setup compels us to use a recursive representation of the objective of the decision-maker.

8 Sequential timing protocol for a constraint formulation

⟨Rconstraint⟩

Section 7 showed how to make penalty problem 5.2 recursive by adopting a sequential timing protocol. Now we show how to make the constraint problem 5.3 recursive. Because the value of the date zero constraint problem depends on the magnitude of the entropy constraint, we add the continuation value of entropy as a state variable. Instead of a value function V that depends only on the state x, we use a value function K that also depends on continuation entropy, denoted r.

8.1 A Bellman equation for a constraint game

Our strategy is to use the link between the value functions for the penalty and constraint problems asserted in claims 5.4 and 5.6, then to deduce from the Bellman equation (32) a partial differential equation that can be interpreted as the Bellman equation for another zero-sum two-player game with additional states and controls. By construction, the new game has a sequential timing protocol and will have the same equilibrium outcome and representation as game (32). Until now we have suppressed the dependence of V on θ in our notation for the value function V. This dependence is central in our analysis and so we now denote it explicitly.

8.2 Another value function

Claim 5.4 showed how to construct the date zero value function for the constraint problem from the penalty problem via Legendre transform. We use this same transform over time to construct the value function K:

$$K(\check{x}, \check{r}) = \max_{\theta \geq 0} V(\check{x}, \theta) - \check{r}\check{z}\theta. \tag{36}$$ legendre5

This new value function is related to $\tilde{K}$ by

$$\tilde{K}(\check{r}) = K(\check{x}, \check{r})$$

provided that $\check{x}$ is equal to the date zero state x_0, $\check{r}$ is used for the initial entropy constraint, and $\check{z} = 1$. We also assume that the Bellman-Isaacs condition is satisfied so that the inverse Legendre transform can be applied:

$$V(\check{x}, \theta) = \min_{\check{r} \geq 0} K(\check{x}, \check{r}) + \check{r}\theta. \tag{37}$$ legendre5inv

When K abd V are related by the Legendre tranforms (36) and (37), their derivatives are closely related, if they exist. In what follows, we presume the smoothness needed to compute derivatives.

The Hamilton-Jacobi-Bellman equation (32) that we derived for V held for each value of θ. In what follows, we consider the consequences of varying the pair $(\hat{x}, \theta)$ as in the construction of V or we consider varying the pair $(\hat{x}, \hat{r})$ as in the construction of K.

$$K_r = \theta \quad \text{or} \quad V_\theta = \hat{r}.$$

For a fixed $\hat{x}$ we can vary $\hat{r}$ by changing θ or conversely we can vary θ by changing $\hat{r}$. To construct a partial differential equation for K from (32) we will compute derivatives with respect to $\hat{r}$ that respect the constraint linking $\hat{r}$ and θ.

For the optimized value of $\check{r}$, we have

$$\delta V = \delta(K + \theta \check{r}) = \delta K + \delta \check{r} K_r, \tag{38}$$ levelK

and

$$\theta \left(\frac{\check{h} \cdot \check{h}}{2} \right) = K_r \left(\frac{\check{h} \cdot \check{h}}{2} \right). \tag{39}$$ evolver

By the implicit function theorem (holding θ fixed):

$$\frac{\partial \check{r}}{\partial x} = - \frac{K_{xr}}{K_{rr}}.$$

Next we compute the derivatives of V that enter the partial differential equation (32) for V:

$$\begin{aligned} V_x &= K_x \\ V_{xx} &= K_{xx} + K_{rx} \frac{\partial \check{r}}{\partial x} \\ &= K_{xx} - \frac{K_{rx} K_{xr}}{K_{rr}}. \end{aligned} \tag{40}$$ firstK

Notice that

$$\begin{aligned} &\tfrac{1}{2} \text{trace} \left[\sigma(\check{c}, \check{x})' V_{xx}(\check{x}) \sigma(\check{c}, \check{x}) \right] = \\ &\min_{\breve{g}} \ \tfrac{1}{2} \text{trace} \left(\begin{bmatrix} \sigma(\check{c}, \check{x})' & \breve{g} \end{bmatrix} \begin{bmatrix} K_{xx}(\check{x}, \check{r}) & K_{xr}(\check{x}, \check{r}) \\ K_{rx}(\check{x}, \check{r}) & K_{rr}(\check{x}, \check{r}) \end{bmatrix} \begin{bmatrix} \sigma(\check{c}, \check{x}) \\ \breve{g}' \end{bmatrix} \right) \end{aligned} \tag{41}$$ Ktrace

where $\breve{g}$ is a column vector with the same dimension d as the Brownian motion. Substituting equations (38), (39), (40), and (41) into the partial differential equation (33) gives:

$$\begin{aligned} \delta K(\check{x}, \check{r}) \;=\; & \max_{\check{c} \in \check{C}} \min_{\check{h}, \breve{g}} U(\check{c}, \check{x}) + \left[\mu(\check{c}, \check{x}) + \sigma(\check{c}, \check{x}) \check{h} \right] \cdot K_x(\check{x}, \check{r}) + \left(\delta \check{r} - \frac{\check{h} \cdot \check{h}}{2} \right) \cdot K_r(\check{x}, \check{r}) \\ & + \frac{1}{2} \text{trace} \left(\begin{bmatrix} \sigma(\check{c}, \check{x})' & \breve{g} \end{bmatrix} \begin{bmatrix} K_{xx}(\check{x}, \check{r}) & K_{xr}(\check{x}, \check{r}) \\ K_{rx}(\check{x}, \check{r}) & K_{rr}(\check{x}, \check{r}) \end{bmatrix} \begin{bmatrix} \sigma(\check{c}, \check{x}) \\ \breve{g}' \end{bmatrix} \right). \end{aligned} \tag{42}$$ HJB_c

The remainder of this section interprets $\check{z} K(\check{x}, \check{r})$ as a value function for a recursive game in which $\theta = \theta^* > \underline{\theta}$ is fixed over time. We have already seen how to characterize the state evolution for the recursive penalty differential game associated with a fixed θ. The first-order condition for the maximization problem on the right side of (36) is

$$\check{r} = V_\theta(\check{x}, \theta^*). \tag{43}$$

cont-entropy

We view this first-order condition as determining $\check{r}$ for a *given* θ^* and $\check{x}$. Then formula (43) implies that the evolution of r is fully determined by the equilibrium evolution of x. We refer to r as *continuation entropy.*

In what follows, we denote the state evolution for the θ^* differential game as:

$$dx_t = \mu^*(x_t, \theta^*)dt + \sigma^*(x_t, \theta^*)dB_t$$

8.3 Continuation entropy

We want to show that r evolves like continuation entropy? Recall formula (28) for the relative entropy of a nonnegative martingale:

$$\mathcal{R}(z) \doteq E \int_0^\infty \exp(-\delta t) z_t \frac{|h_t|^2}{2} dt.$$

Define a date t conditional counterpart as follows:

$$\mathcal{R}_t(z) = E\left[\int_0^\infty \exp(-\delta u) \left(\frac{z_{t+u}}{z_t} \right) \frac{|h_{t+u}|^2}{2} du \Big| \mathcal{F}_t \right]$$

provided that $z_t > 0$ and define $\mathcal{R}_t(z)$ to be zero otherwise. This family of random variables induces the following recursion for $\epsilon > 0$:

$$z_t \mathcal{R}_t(z) = \exp(-\delta\epsilon) E\left[z_{t+\epsilon} \mathcal{R}_{t+\epsilon}(z) \Big| \mathcal{F}_t \right] + E\left[\int_0^\epsilon \exp(-\delta u) z_{t+u} \frac{|h_{t+u}|^2}{2} du \Big| \mathcal{F}_t \right].$$

Since $z_t \mathcal{R}_t(z)$ is in the form of a risk neutral value of an asset with future *dividend* $z_{t+u} \frac{h_{t+u} \cdot h_{t+u}}{2}$, its local mean or drift has the familiar formula:

$$\delta z_t \mathcal{R}_t(z) - z_t \frac{|h_t|^2}{2}.$$

To defend an interpretation of r_t as continuation entropy, we need to verify that this drift restriction is satisfied for $r_t = \mathcal{R}_t(z)$. Write the evolution for r_t as:

$$dr_t = \mu_r(x_t)dt + \sigma_r(x_t) \cdot dB_t,$$

and recall that

$$dz_t = z_t h_t \cdot dB_t.$$

Using Ito's formula for the drift of $z_t r_t$, drift restriction (44) that we seek to verify is:

$$\check{z}\mu_r(\check{x}) + \check{z}\sigma_r(\check{x}) \cdot \check{h} = \delta \check{z}\check{r} - \check{z}\frac{|\check{h}|^2}{2}. \tag{44}$$

driftcheck

Given formula (43) and Ito's differential formula for a smooth function of a diffusion process, we have

$$\mu_r(\check{x}) = V_{\theta x}(\check{x}, \theta^*) \cdot \mu^*(\check{x}, \theta^*) + \frac{1}{2}\text{trace}\left[\sigma(\check{c}, \check{x})' V_{\theta xx}(\check{x}) \sigma(\check{c}, \check{x})\right]$$

and

$$\sigma_r(\check{x}) = V_{\theta x}(\check{x}, \theta^*) \sigma^*(\check{x}, \theta^*).$$

Recall that the worst case h_t is given by

$$h_t = -\frac{1}{\theta^*} \sigma^*(x_t, \theta^*)' V_x(x_t, \theta^*)$$

and thus

$$\frac{|h_t|^2}{2} = \left(\frac{1}{2\theta^{*2}}\right) V_x(\check{x})' \sigma(\check{c}, \check{x}) \sigma(\check{c}, \check{x})' V_x(\check{x}).$$

Restriction (44) can be verified by substituting our formulas for r_t, h_t, μ_r and σ_r. The resulting equation is equivalent to that obtained by differentiating the Hamilton-Jacobi-Bellman equation (35) with respect to θ, justifying our interpretation of r_t as a continuation entropy.

8.4 Minimizing continuation entropy

Having defended a specific construction of continuation entropy that supports a constant value of θ, we now describe a differential game that makes entropy an endogenous state variable. To formulate that game, we consider the inverse Legendre transform (37) from which we construct V from K by minimizing $\check{r}$. In the recursive version of the constraint game, the state variable r_t is the continuation entropy that at t remains available to allocate across states at future dates. At date t, continuation entropy is allocated via the minimization suggested by the inverse Legendre transform. We restrict the minimizing player to allocate future r_t across states that can be realized with positive probability, conditional on date t information.

8.4.1 Two state example

Before presenting the continuous-time formulation, consider a two-period example. Suppose that two states can be realized at date $t+1$, namely ω_1 and ω_2. Each state has probability one-half under an approximating model. The minimizing agent distorts these probabilities by assigning probability p_t to state ω_1. The contribution to entropy coming from the distortion of the probabilities is the discrete state analogue of $\int \log\left(\frac{dq_t}{dq_t^0}\right) dq_t$, namely,

$$I(p_t) = p_t \log p_t + (1 - p_t)\log(1 - p_t) + \log 2.$$

The minimizing player also chooses continuation entropies for each of the two states that can be realized next period. Continuation entropies are discounted and averaged according to the distorted probabilities, so that we have:

$$r_t = I(p_t) + \exp(-\delta)\left[p_t r_{t+1}(\omega_1) + (1-p_t) r_{t+1}(\omega_2)\right]. \tag{45}$$ dis:constrain

Let U_t denote the current period utility (we abstract from maximization), and let $V_{t+1}(\omega, \theta)$ denote the next period value given state ω. This function is concave in θ. Construct V_t via backward induction:

$$\begin{aligned} V_t(\theta) &= \min_{0\le p_{t+1}\le 1} U_t + \theta I_t(p_t) \\ &\quad + \exp(-\delta)\left[p_t V_{t+1}(\omega_1,\theta) + (1-p_t) V_{t+1}(\omega_2,\theta)\right] \end{aligned} \tag{46}$$ bell:two

Compute the Legendre transforms:

$$\begin{aligned} K_t(\check{r}) &= \max_{\theta\ge 0} V_t(\theta) - \theta\check{r} \\ K_{t+1}(\check{r},\omega) &= \max_{\theta\ge 0} V_{t+1}(\theta) - \theta\check{r} \end{aligned}$$

for $\omega = \omega_1, \omega_2$. Given θ^*, let r_t be the solution to the inverse Legendre transform:

$$V_t(\theta^*) = \min_{\check{r}\ge 0} K_t(\check{r}) + \theta^*\check{r}.$$

Similarly, let $r_{t+1}(\omega)$ be the solution to

$$V_{t+1}(\omega,\theta^*) = \min_{\check{r}\ge 0} K_{t+1}(\omega,\check{r}) + \theta^*\check{r}.$$

Substitute the inverse Legendre transforms into the simplified Bellman equation (46):

$$\begin{aligned} V_t(\theta^*) &= \min_{0\le p_t\le 1} U_t + \theta^* I_t(p_t) \\ &\quad + \exp(-\delta)\left(p_t\left[\min_{\check{r}_1\ge 0} K_{t+1}(\omega_1,\check{r}_1) + \theta\check{r}_1\right] + (1-p_t)\left[\min_{\check{r}_2\ge 0} K_{t+1}(\omega_2,\check{r}_2) + \theta^*\check{r}_2\right]\right) \\ &= \min_{0\le p_t\le 1,\check{r}_1\ge 0,\check{r}_2\ge 0} U_t + \theta^*\left[I_t(p_t) + p_t\check{r}_1 + (1-p_t)\check{r}_2\right] \\ &\quad + \exp(-\delta)\left[p_t K_{t+1}(\omega_1,\check{r}_1) + (1-p_t)K_{t+1}(\omega_2,\check{r}_2)\right]. \end{aligned}$$

Thus

$$\begin{aligned} K_t(r_t) &= V_t(\theta^*) - \theta^* r_t \\ &= \min_{0\le p_t\le 1,\check{r}_1\ge 0,\check{r}_2\ge 0}\ \max_{\theta\ge 0} U_t + \theta\left[I_t(p_t) + p_t\check{r}_1 + (1-p_t)\check{r}_2 - r_t\right] \\ &\quad + \exp(-\delta)\left[p_t K_{t+1}(\omega_1,\check{r}_1) + (1-p_t)K_{t+1}(\omega_2,\check{r}_2)\right]. \end{aligned}$$

Since the solution is $\theta = \theta^* > 0$, at this value of θ the entropy constraint (45) must be satisfied and

$$K_t(r_t) = \min_{0\le p_t\le 1,\check{r}_1\ge 0,\check{r}_2\ge 0} U_t + \exp(-\delta)\left[p_t K_{t+1}(\omega_1,\check{r}_1) + (1-p_t)K_{t+1}(\omega_2,\check{r}_2)\right].$$

By construction, the solution for $\check{r}_j$ is $r_{t+1}(\omega_j)$ defined earlier. The recursive implementation presumes that the continuation entropies $r_{t+1}(\omega_j)$ are chosen at date t prior to the realization of ω.

When we include the control c_t, this construction requires that we can freely change orders of maximization and minimization as in our previous analysis.

8.4.2 Continuous-time formulation

In a continuous-time formulation, we allocate the stochastic differential of entropy subject to the constraint that the current entropy is r_t. The increment to r is determined via the stochastic differential equation:[15]

$$dr_t = \left(\delta r_t - \frac{|h_t|^2}{2} - g_t \cdot h_t \right) dt + g_t \cdot dB_t.$$

This evolution for r implies that

$$d(z_t r_t) = \left(\delta z_t r_t - z_t \frac{|h_t|^2}{2} \right) dt + z_t (r_t h_t + g_t) dB_t$$

which has the requisite drift to interpret r_t as continuation entropy.

The minimizing agent not only picks h_t but also chooses g_t to allocate entropy over the next instant. The process g thus becomes a control vector for allocating continuation entropy across the various future states. In formulating the continuous-time game, we thus add a state r_t and a control g_t to the states and controls. With these added states, the differential game has a value function $\hat{z}K(\hat{x}, \hat{r})$, where K satisfies the Hamilton-Jacobi-Bellman equation (42).

We have deduced this new partial differential equation partly to help us understand senses in which the constrained problem is or is not time consistent. Since r_t evolves as an exact function of x_t, it is more efficient to compute V and to use this value function to infer the optimal control law and the implied state evolution. In the next section, however, we use the recursive constraint formulation to address some interesting issues raised by Epstein and Schneider (2001).

9 A recursive multiple priors formulation

⟨larrymartin⟩ Taking continuation entropy as a state variable is a convenient way to restrict the models entertained at time t by the minimizing player in the recursive version of constraint game. Suppose instead that at date t the decision maker retains the date zero family of prior distributions without imposing additional restrictions or freezing a state variable like continuation entropy. That would allow the minimizing decision maker at date t to reassign probabilities

[15] The process is stopped if r_t hits the zero boundary. Once zero is hit, the continuation entropy remains at zero. In many circumstances, the zero boundary will never be hit.

of events that have already been realized and events that cannot possibly be realized given current information. The minimizing decision maker would take advantage of that opportunity. Therefore, the minimizing player would alter worst-case probability distribution at date t in a way that makes the specification of prior distributions of section 5 dynamically inconsistent in a sense formalized by Epstein and Schneider (2001). They characterize families of prior distributions that satisfy a rectangularity criterion that shields the decision maker from "dynamic inconsistency" in their sense. In this section, we briefly discuss how Epstein and Schneider's notion of dynamic inconsistency would apply to our setting.

Consider the martingale formulation of the date zero entropy constraint:

$$E \int_0^\infty \exp(-\delta u) z_u \frac{|h_u|^2}{2} \leq \eta \tag{47}$$

mart-constrai

where

$$dz_t = z_t h_t \cdot dB_t.$$

The component of entropy that contains our date t decision-maker is:

$$r_t = \frac{1}{z_t} E \left(\int_0^\infty z_{t+u} \frac{|h_{t+u}|^2}{2} | \mathcal{F}_t \right)$$

in states in which $z_t > 0$. We rewrite (47) as:

$$E \int_0^t \exp(-\delta u) z_u \frac{|h_u|^2}{2} du + \exp(-\delta t) E z_t r_t \leq \eta.$$

To illuminate the nature of dynamic inconsistency, we begin by noting that the time 0 constraint imposes essentially no restriction on r_t. Consider a date t event that has probability strictly less than one conditioned on date zero information. Let y be a random variable that is equal to zero on the event and equal to the reciprocal of the probability on the complement of the event. Thus y is a nonnegative, bounded random variable with expectation equal to unity. Construct a $z_u = E(y|\mathcal{F}_u)$. Then z is a bounded nonnegative martingale with finite entropy and $z_u = y$ for $u \geq t$. In particular z_t is zero on the date t event used to construct y. By shrinking the date t event to have arbitrarily small probability, we can bring the bound arbitrarily close to unity and entropy arbitrarily close to zero. Thus for date t events with sufficiently small probability, the entropy constraint can be satisfied without restricting the magnitude of r_t on these events. This exercise isolates a justification for using continuation entropy as a state variable inherited at date t: fixing it eliminates any gains from readjusting distortions of probabilities in previous time periods.

If we insist on withdrawing an endogenous state variable like r_t, dynamic consistency can still be obtained by imposing restrictions on h_t for alternative dates and states. For instance, we could impose prior restrictions in the separable form

$$\frac{|h_t|^2}{2} \leq f_t$$

for each event realization and date t. Such a restriction is *rectangular* in the sense of Epstein and Schneider (2001). To preserve a subjective notion of prior distributions, Epstein and Schneider (2001) advocate enlarging an original set of priors to be rectangular to the least extent possible. They suggest this approach in conjunction with entropy measures of the type used here and other possible specifications. An f_t specified on any event that occurs with probability less than one is essentially unrestricted by the date zero entropy constraint, however. In continuous time this follows because zero measure is assigned to any calendar date, but it also carries over to discrete time for basically the same reason that continuation entropy remains unrestricted if we can adjust earlier distortions. For our application this way of achieving a rectangular specification through the mechanism suggested by Epstein and Schneider (2001) fails to restrict prior distributions in an interesting way.[16]

There is an alternative way to make the priors rectangular that has trivial consequences for our analysis. The ability to exchange maximization and minimization is central to this construction. From section 5, recall that

$$\tilde{K}(\check{r}) = \max_{\theta>0} \tilde{V}(\theta) - \theta\check{r}.$$

We now rewrite the inner problem on the right side for a fixed θ. Take the Bellman-Isaacs condition

$$zV(x) = \min_{h\in H} \max_{c\in C} \max_{c\in C} E \int_0^\infty \exp(-\delta t)\left[z_t U(c_t, x_t) + \theta z_t \frac{|h_t|^2}{2}\right]$$

with the evolution equations

$$\begin{aligned} dx_t &= \mu(c_t, x_t)dt + \sigma(c_t, x_t)dB_t \\ dz_t &= z_t h_t \cdot dB_t. \end{aligned} \tag{48}$$

mart-evolve

Decompose the entropy constraint as:

$$\eta = E\int_0^\infty \exp(-\delta t) z_t f_t dt$$

where

$$f_t = \frac{|h_t|^2}{2}.$$

Rewrite the objective of the optimization problem as

$$\min_{f\in F} \min_{h\in H, \frac{|h_t|^2}{2}\le f_t} \max_{c\in C} E\int_0^\infty \exp(-\delta t)\left[z_t U(c_t, x_t) + \theta z_t g_t\right]$$

subject to (48). In this formulation, F is the set of progressively measurable scalar processes that are nonnegative. We entertain the inequality

$$\frac{|h_t|^2}{2} \le f_t$$

[16] While Epstein and Schneider (2001) advocate rectangularization even for entropy-based constraints, they do not claim that it always gives rise to interesting restrictions.

but in fact this constraint will always bind for the *a priori* optimized choice of f. The inner problem can now be written as:

$$\min_{h \in H, \frac{|h_t|^2}{2} \leq f_t} \max_{c \in C} E \int_0^\infty \exp(-\delta t) z_t U(c_t, x_t)$$

subject to (48). Provided that we can change orders of the min and max, this inner problem will have a rectangular specification of alternative models and be dynamically consistent in the sense of Epstein and Schneider (2001).

Although this construction avoids introducing continuation entropy as an endogenous state variable, it assumes a commitment to a process f that is computed *ex ante* by solving what is essentially a static optimization problem. That is, f is chosen by exploring its consequences for a dynamic implementation of the form envisioned by Epstein and Schneider (2001) and is not simply part of the exogenously *ex ante* given set of beliefs of the decision maker.[17] While their concern about dynamic consistency leads Epstein and Schneider to express doubts about commitments to a constraint based on continuation entropy, they do not question what could lead a decision-maker to commit to a particular rectangular set of beliefs as reflected in a specification of f, just as Bayesian decision theory does not question a commitment to a specfic prior distribution.

If the multiple priors truly are a subjective statement of beliefs, then one is hard pressed to dismiss these beliefs on the grounds of dynamic inconsistency and rectangularity simply becomes a useful statement of what might or might not be argued is a weaker form of commitment that suffices to make the decision problem recursive. But making general classes of models rectangular through enlargement reduces the content of the prior beliefs. In our context this enlargement is immense.

The reservations that we have expressed about the substantive importance of rectangularity notwithstanding, we agree that Epstein and Schneider's discussion of dynamic consistency opens up a useful discussion of the alternative possible forms of commitment that allow us to create dynamic models with multiple priors.

10 Concluding remarks

⟨Concl⟩

Empirical studies in macroeconomics and finance typically assume a unique and explicitly specified dynamic statistical model. Concerns about model misspecification admit that one of a set of alternative models might instead govern the data. But how should one specify those alternative models? With one parameter that measures the size of the set, robust control theory parsimoniously stipulates a set of alternative models with rich alternative dynamics. The theory leaves those models only vaguely specified and obtains them by perturbing the decision maker's approximating model to let shocks feed back on state variables arbitrarily. Among other possibilities, this allows the approximating model to miss the se-

[17] Notice that the Bayesian interpretation is also a trivial special case of a recursive multiple priors model.

rial correlation of exogenous variables and the dynamics of how those exogenous variables impinge on endogenous state variables.

We have delineated the formal connections that exist between various formulations of robust control theory and the maxmin of Gilboa and Schmeidler (1989). Their theory deduces a set of models from a decision maker's underlying preferences over risky outcomes. In their theory, none of the decision maker's models has the special status that the approximating model has in robust control theory. To put Gilboa and Schmeidler's theory to work, an applied economist would have to impute a set of models to the decision makers in his model (unlike the situation in rational expectations models, where the decision maker's model would be an equilibrium outcome). A practical attraction of robust control theory is the way it allows an economist to take a single approximating model and from it to manufacture a set of models that express a decision maker's ambiguity. Hansen and Sargent (2003) exploit this feature of robust control to construct a multiple agent model in which a common approximating model plays the role that an equilibrium common model does in a rational expectations model.

We have used a particular notion of discounted entropy as a statistical measure of the discrepancy between models. It our decision maker's attention to models that are absolutely continuous with respect to his approximating model over finite intervals, but not absolutely continuous with respect to it over an infinite interval. This specification keeps the decision maker concerned about models that can be difficult to distinguish from the approximating model from a continuous record of observations on the state vector of a finite length. Via statistical detection error probabilities, Anderson, Hansen, and Sargent (2003) show how the multiplier parameter or the constraint parameter in the robust control problems can be used to identify a set of perturbed models that are difficult to distinguish statistically from the approximating model in light of a continuous record of finite length T of observations on x_t.

⟨cast⟩

A Cast of characters

This appendix sets out the following list of objects and conventions that make repeated appearances in our analysis.

1. Probability spaces

⟨brown-fixed⟩

   (a) A probability space associated with a Brownian motion B that is used to define an approximating model and a set of alternative models.

   (b) A probability space over continuous functions of time induced by history of the Brownian motion B in part 1a and used to define an approximating model.

⟨uced-flexible⟩

   (c) A set of alternative probability distributions induced by B and used to define a set alternative models.

2. Ordinary (single-agent) control problems

(a) A benchmark optimal control problem defined on space 1a.

(b) A benchmark decision problem defined on the probability space induced by B.

(c) A risk-sensitive problem defined on space 1a.

(d) Alternative Bayesian (benchmark problems) defined on the spaces in 1c.

3. Representations of alternative models

   (a) As nonnegative martingales with unit expectation the probability space 1a.

   (b) As alternative induced distributions as in 1c.

4. Restrictions on sets of alternative models

   (a) An implicit restriction embedded in a nonnegative penalty parameter θ.

   (b) A constraint on relative entropy, a measure of model discrepancy.

5. Representations of relative entropy

   (a) Time 0 (static): discounted expected log likelihood ratio of an approximating model q^0 to an alternative model q drawn from the set 1c.

   (b) Time 0 (static): a function of a martingale defined on the probability space 1a.

   (c) Recursive: as a solution of either of a differential equations defined in terms of B.

6. Timing protocols for zero-sum two-player games

   (a) Exchange of order of choice for maximizing and minimizing players.

   (b) Under two-sided commitment at $t = 0$, both players choose processes for all time $t \geq 0$.

   (c) With lack of commitment on two sides, both players choose sequentially.

B Discounted entropy

appen:entropy⟩ Let Q be the set of all distributions that are absolutely continuous with respect to q^0 over finite intervals. This set is convex. For $q \in Q$, let

$$\tilde{\mathcal{R}}(q) \doteq \delta \int_0^\infty \exp(-\delta t) \log \left(\frac{dq_t}{dq_t^o} \right) dq_t dt,$$

which may be infinite for some $q \in Q$.

append:claim1⟩ **Claim B.1.** *$\tilde{\mathcal{R}}$ is convex on Q.*

Proof. Since $q \in Q$ is absolutely continuous with respect to q^0 over finite intervals, we can construct likelihood ratios for finite histories at any calendar date t. Form $\tilde{\Omega} = \Omega^* \times \mathbb{R}^+$ where $\mathbb{R}^+$ is the nonnegative real line. Form the corresponding sigma algebra $\tilde{\mathcal{F}}$ as the smallest sigma algebra containing $\mathcal{F}_t^* \otimes \mathcal{B}_t$ for any t where $\mathcal{B}_t$ is the collection of Borel sets in $[0, t]$; and form $\tilde{q}$ as the product measure q with an exponential distribution with density $\delta \exp(-\delta t)$ for any $q \in Q$. Notice that $\tilde{q}$ is a probability distribution and $\tilde{\mathcal{R}}(q)$ is the relative entropy of $\tilde{q}$ with respect to $\tilde{q}^o$:

$$\tilde{\mathcal{R}}(q) = \int \log \left(\frac{d\tilde{q}^o}{d\tilde{q}} \right) d\tilde{q}.$$

Form two measures $\tilde{q}^1$ and $\tilde{q}^2$ as the product of q^1 and q^2 with an exponential distribution with parameter δ. Then a convex combination of $\tilde{q}^1$ and $\tilde{q}^2$ is given by the product of the corresponding convex combination of q^1 and q^2 with the same exponential distribution. Relative entropy is well known to be convex in the probability measure $\tilde{q}$, and hence $\tilde{R}$ is convex in q. (Add a reference.) □

Recall that associated with any probability measure q that is absolutely continuous with respect to q^o over finite intervals is a nonnegative martingale z defined on $(\Omega, \mathcal{F}, P)$ with a unit expectation. This martingale satisfies the integral equation:

$$z_t = 1 + \int_0^t z_u h_u dB_u. \tag{49}$$ append:int

append:claim2⟩ **Claim B.2.** *Suppose that q_t is absolutely continuous with respect to q_t^0 for all $0 < t < \infty$. Let z be the corresponding nonnegative martingale on $(\Omega, \mathcal{F}, P)$. Then*

$$E z_t \mathbf{1}_{\left\{\int_0^t |h_s|^2 ds < \infty\right\}} = 1.$$

Moreover,

$$\int \log \frac{dq_t}{dq_t^0} dq_t = \frac{1}{2} E \int_0^t z_s |h_s|^2 ds.$$

Proof. Consider first the claim that

$$E z_t \mathbf{1}_{\left\{\int_0^t |h_s|^2 ds < \infty\right\}} = 1,$$

The martingale z satisfies the stochastic differential equation:

$$dz_t = z_t h_t dB_t$$

with initial condition $z_0 = 1$. Construct an increasing sequence of stopping times $\{\tau_n : n \geq 1\}$ where $\tau_n \doteq \inf\{t : z_t = \frac{1}{n}\}$ and let $\tau = \lim_n \tau_n$. The limiting stopping time can be infinite. Then $z_t = 0$ for $t \geq \tau$ and

$$z_t = z_{t \wedge \tau}$$

Form:

$$z_t^n = z_{t\wedge\tau_n}$$

which is nonnegative martingale satisfying:

$$dz_t^n = z_t^n h_t^n dB_t$$

where $h_t^n = h_t$ if $0 < t < \tau_n$ and $h_t^n = 0$ if $t \geq \tau_n$. Then

$$P\left\{\int_0^t |h_s^n|^2 (z_s^n)^2 < \infty\right\} = 1$$

and hence

$$P\left\{\int_0^t |h_s^n|^2 ds < \infty\right\} = P\left\{\int_0^{t\wedge\tau_n} |h_s|^2 ds < \infty\right\} = 1.$$

Taking limits as n gets large,

$$P\left\{\int_0^{t\wedge\tau} |h_s|^2 ds < \infty\right\} = 1.$$

While it is possible that $\tau < \infty$ with positive P probability, as argued by Kabanov, Lipcer, and Sirjaev (1979)

$$\int z_t \mathbf{1}_{\{\tau<\infty\}} dP = \int_{\{z_t=0,\ t<\infty\}} z_t dP = 0.$$

Therefore,

$$Ez_t \mathbf{1}_{\left\{\int_0^t |h_s|^2 ds < \infty\right\}} = Ez_t \mathbf{1}_{\left\{\int_0^{t\wedge\tau} |h_s|^2 ds<\infty, \tau=\infty\right\}} + Ez_t \mathbf{1}_{\left\{\int_0^t |h_s|^2 ds<\infty, \tau<\infty\right\}} = 1.$$

Consider next the claim that

$$\int \log \frac{dq_t}{dq_t^0} dq_t = E\int_0^t z_s |h_s|^2 ds.$$

We first suppose that

$$E\int_0^t z_s |h_s|^2 ds < \infty. \tag{50}$$

finite-energy

We will subsequently show that this condition is satisfied when $\tilde{\mathcal{R}}(q) < \infty$. Use the martingale z to construct a new probability measure $\tilde{P}$ on $(\Omega, \mathcal{F})$. Then from the Girsanov Theorem [see Theorem 6.2 of Liptser and Shiryaev (2000)]

$$\tilde{B}_t = B_t - \int_0^t h_s ds$$

is a Brownian motion with respect to the filtration $\{\mathcal{F}_t : t \geq 0\}$. Moreover,

$$\tilde{E}\int_0^t |h_s|^2 ds = E\int_0^t z_s |h_s|^2.$$

Write

$$\log z_t = \int_0^t h_s \cdot dB_s - \frac{1}{2}\int_0^t |h_s|^2 ds = \int_0^t h_s \cdot d\tilde{B}_s + \frac{1}{2}\int_0^t |h_s|^2 ds.$$

which is well defined under the $\tilde{P}$ probability. Moreover,

$$\tilde{E}\int_0^t h_s \cdot d\tilde{B}_s = 0$$

and hence

$$\tilde{E}\log z_t = \frac{1}{2}\tilde{E}\int_0^t |h_s|^2 ds = \frac{1}{2}E\int_0^t z_s|h_s|^2 ds,$$

which is the desired equality. In particular, we have proved that $\int \log \frac{dq_t}{dq_t^0} dq_t$ is finite.

Next we suppose that

$$\int \log \frac{dq_t}{dq_t^0} dq_t < \infty,$$

which will hold when $\tilde{\mathcal{R}}(q) < \infty$. Then Lemma 2.6 from Föllmer (1985) insures that

$$\frac{1}{2}\tilde{E}\int_0^t |h_s|^2 ds \leq \int \log \frac{dq_t}{dq_t^0} dq_t.$$

Föllmer's result is directly applicable because $\int \log \frac{dq_t}{dq_t^0} dq_t$ is the same as the relative entropy of $\tilde{P}_t$ with respect to P_t where $\tilde{P}_t$ is the restriction of $\tilde{P}$ to events in $\mathcal{F}_t$ and P_t is defined similarly. As a consequence, (50) is satisfied and the desired equality follows from our previous argument.

Finally, notice that $\frac{1}{2}\tilde{E}\int_0^t |h_s|^2 ds$ is infinite if, and only if $\int \log \frac{dq_t}{dq_t^0} dq_t$ is infinite. □

Claim B.3. *For $q \in Q$, let z be the nonnegative martingale associated with q and let h be the progressively measurable process satisfying (49). Then*

$$\tilde{\mathcal{R}}(q) = \frac{1}{2}E\left[\int_0^\infty \exp(-\delta t) z_t |h_t|^2 dt\right]$$

Proof. The conclusion follows from:

$$\begin{aligned}
\tilde{\mathcal{R}}(q) &= \delta \int_0^\infty \exp(-\delta t) \int \log\left(\frac{dq_t}{dq_t^0}\right) dq_t dt \\
&= \frac{\delta}{2}E\left[\int_0^\infty \exp(-\delta t)\int_0^t z_u|h_u|^2 du dt\right] \\
&= \frac{1}{2}E\left[\int_0^\infty \exp(-\delta t) z_t|h_t|^2 dt\right]
\end{aligned}$$

where the second equality follows from B.2 and the third from integrating by parts. □

This justifies our definition of entropy for nonnegative martingales:

$$\mathcal{R}(z) = \frac{1}{2}E\left[\int_0^\infty \exp(-\delta t) z_t |h_t|^2 dt\right].$$

C Absolute continuity of solutions

⟨appen-mart⟩ In this appendix we show how to verify that the solution for z from the martingale robust control problem is in a fact a martingale and not just a local martingale. Our approach to studying absolute continuity and verifying that the Markov perfect equilibrium z is a martingale differs from the perhaps more familiar use of a Novikov or Kazamaki condition.[18]

Consider two distinct stochastic differential equations. One is the Markov solution to the penalty robust control problem.

$$\begin{aligned} dx_t^* &= \mu^*(x_t^*) + \sigma^*(x_t^*) dB_t \\ dz_t^* &= z_t^* \alpha_h(x_t^*) dB_t. \end{aligned} \tag{51}$$

zevol

where $\mu^*(\check{x}) = \mu(\alpha_c(\check{x}), \check{x}), \sigma^*(\check{x}) = \sigma(\alpha_c(\check{x}), \check{x})$ and where α_c and α_h are the solutions from the penalty robust control problem. Notice that the equation for the evolution of x_t^* is autonomous (it does not depend on z_t^*). Let a strong solution to this equation system be:

$$x_t^* = \Phi_t^*(B).$$

Consider a second stochastic differential equation:

$$d\hat{x}_t = \mu^*(\hat{x}_t)dt + \sigma^*(\hat{x}_t)\left[\alpha_h(\hat{x}_t) + d\hat{B}_t\right] \tag{52}$$

hatevol

In verifying that this state equation has a solution, we are free to examine weak solutions provided that $\hat{\mathcal{F}}_t$ is generated by current and past $\hat{x}_t$ and $\hat{B}$ does not generate a larger filtration than $\hat{x}$.

The equilibrium outcomes x^* and $\hat{x}$ for the two stochastic differential equations thus induce two distributions for x. We next study how these distributions are related. We will discuss how to check for absolute continuity along finite intervals for induced distributions associated with these models. When the models satisfy absolute continuity over finite intervals, it will automatically follow that the equilibrium process z^* is a martingale.

C.1 Comparing models of B

We propose the following method to transform a strong solution to (51) into a possibly weak solution to (52). Begin with a Brownian motion $\hat{B}$ defined a probability space with probability measure $\hat{P}$. Consider the recursive solution:

$$\begin{aligned} \hat{x}_t &= \Phi_t^*(B) \\ B_t &= \hat{B}_t + \int_0^t \alpha_h(\hat{x}_u) du. \end{aligned}$$

[18] We construct two well defined Markov processes and verify absolute continuity. Application of the Novikov or Kazamaki conditions entails imposing extra moment conditions on the objects used to construct the local martingale z.

We look for solutions in which $\hat{\mathcal{F}}_t$ is generated by current and past values of B (not $\hat{B}$). We call this a recursion because B is itself constructed from past values of B and $\hat{B}$. The stochastic differential equation associated with this recursion is (52).

To establish the absolute continuity of the distribution induced by B with respect to Weiner measure q^0 it suffices to verify that for each t

$$\hat{E}\int_0^t |\alpha_h(\hat{x}_u)|^2 du < \infty \tag{53}$$

?absoluteinq?

and hence

$$\hat{P}\left\{\int_0^t |\alpha_h(\hat{x}_u)|^2 du < \infty\right\} = 1. \tag{54}$$

ineq:abscont

It follows from Theorem 7.5 of Liptser and Shiryaev (2000) that the probability distribution induced by B under the solution to the perturbed problem is absolutely continuous with respect to Wiener measure q^0. To explore directly the weaker relation (54) further, recall that

$$\alpha_h(\check{x}) = -\frac{1}{\theta}\sigma^*(\check{x})' V_x(\check{x}).$$

Provided that σ^* and V_x are continuous in $\check{x}$ and that x does not explode in finite time, this relation follows immediately.

C.2 Comparing generators

Another strategy for checking absolute continuity is to follow the approach of Kunita (1969), who provides characterizations of absolute continuity and equivalence of Markov models through restrictions on the generators of the processes. Since the models for x^* and $\hat{x}$ are Markov diffusion processes, we can apply these characterizations provided that we include B as part of the state vector. Abstracting from boundary behavior, Kunita (1969) requires a common diffusion matrix, which can be singular. The differences in the drift vector are restricted to be in the range of the common diffusion matrix. These restrictions are satisfied in our application.

C.3 Verifying z^* is a martingale

We apply our demonstration of absolute continuity to reconsider the super martingale z^*. Let κ_t denote the Radon-Nikodym derivative for the two models of B. Conjecture that

$$z_t^* = \kappa_t(B).$$

By construction, z^* is a nonnegative martingale defined on $(\Omega, \mathcal{F}, P)$. Moreover, it is the unique solution to the stochastic differential equation (51) subject to the initial condition $z_0^* = 1$. See Theorem 7.6 of Liptser and Shiryaev (2000).

D Three ways to verify the Bellman-Isaacs condition

BellmanIsaacs⟩ This appendix describes three alternative conditions that are sufficient to verify the Bellman-Isaacs condition embraced in Assumption 7.1.[19] The ability to exchange orders of extremization in the recursive game implies that the orders of extremization can also be exchanged in the static game, as required in Assumption 5.5. As we shall now see, the exchange of order of extremization asserted in Assumption 7.1 can often be verified without knowing the value function S.

D.1 No binding inequality restrictions

Suppose that there are no binding inequality restrictions on c. Then a justification for Assumption 7.1 can emerge from the first-order conditions for $\check{c}$ and $\check{h}$. Define

$$\begin{aligned} \chi(\check{c}, \check{h}, \check{x}) \doteq U(\check{c}, \check{x}) + \frac{\theta}{2}\check{h} \cdot \check{h} \quad &+ \quad \left[\mu(\check{c}, \check{x}) + \sigma(\check{c}, \check{x})\check{h}\right] \cdot S_x(\check{x}) \\ &+ \quad \frac{1}{2}\text{trace}\left[\sigma(\check{c}, \check{x})' S_{xx}(\check{x}) \sigma(\check{c}, \check{x})\right], \end{aligned} \tag{55}$$ {?}

and suppose that χ is continuously differentiable in $\check{c}$. First, find a Markov perfect equilibrium by solving:

$$\begin{aligned} \frac{\partial \chi}{\partial c}(\check{c}^*, \check{h}^*, \check{x}) &= 0 \\ \frac{\partial \chi}{\partial h}(\check{c}^*, \check{h}^*, \check{x}) &= 0. \end{aligned}$$

In particular, the first-order conditions for $\check{h}$ are:

$$\frac{\partial \chi}{\partial h}(\check{c}^*, \check{h}^*, \check{x}) = \theta \check{h}^* + \sigma(\check{c}^*, \check{x})' S_x(\check{x}) = 0.$$

If a unique solution exists and if it suffices for extremization, the Bellman-Isaacs condition is satisfied. This follows from the "chain rule." Thus, suppose that the minimizing player goes first and computes $\check{h}$ as a function of $\check{x}$ and $\check{c}$:

$$\check{h}^* = -\frac{1}{\theta}\sigma(\check{c}, \check{x})' S_x(\check{x}) \tag{56}$$ react

Then the first-order conditions for the max player selecting $\check{c}$ as a function of $\check{x}$ are:

$$\frac{\partial \chi}{\partial c} + \frac{\partial h'}{\partial c}\frac{\partial \chi}{\partial h} = 0$$

where $\frac{\partial h}{\partial c}$ can be computed from the reaction function (56). Notice that the first-order conditions for the maximizing player are satisfied at the Markov perfect equilibrium. A similar argument can be made if the maximizing player chooses first.

[19] Fleming and Souganidis (1989) show that the freedom to exchange orders of maximization and minimization guarantees that equilibria of the static (i.e., choices under mutual commitment at date 0) and the recursive games (i.e., sequential choices by both agents) coincide.

D.2 Separability

Consider next the case in which σ does not depend on the control. In this case the decision problems for $\check{c}$ and $\check{h}$ separate. For instance, from (56), we see that $\check{h}$ does not react to $\check{c}$ in the minimization of $\check{h}$ conditioned on $\check{c}$. Even with binding constraints on $\check{c}$, the Bellman-Isaacs condition (Assumption 7.1) is satisfied, provided that a solution exists for $\check{c}$.

D.3 Convexity

A third approach that uses results of Fan (1952) and Fan (1953) is based on the global shape properties of the objective. When we can reduce the choice set C to be a compact subset of a linear space, Fan (1952) can apply. Fan (1952) also requires that the set of conditional minimizers and maximizers be convex. We know from formula (56) that the minimizers of $\chi(\check{c}, \cdot, \check{x})$ form a singleton set, which is convex for each $\check{c}$ and $\check{x}$.[20] Suppose also that the set of maximizers of $\chi(\cdot, \check{h}, \check{x})$ is non-empty and convex for each $\check{h}$ and $\check{x}$.[21] Then again the Bellman-Isaacs condition (Assumption 7.1) is satisfied. Finally Fan (1953) does not require that the set $\check{C}$ be a subset of a linear space, but instead requires that $\chi(\cdot, \check{h}, \check{x})$ be concave. By relaxing the linear space structure we can achieve compactness by adding points (say the point ∞) to the control set, provided that we can extend $\chi(\cdot, \check{h}, \check{x})$ to be upper semi-continuous. The extended control space must be a compact Hausdorff space. Provided that the additional points are not attained in optimization, we can apply Fan (1953) to verify Assumption 7.1.[22]

E Recursive version of Stackelberg game and a Bayesian problem

⟨appen:bayes⟩

E.1 Recursive version of a Stackelberg game

We first change the timing protocol for decision-making, moving from the Markov perfect equilibrium that gives rise to a value function V to a date zero Stackelberg equilibrium with value function N. In the matrix manipulations that follow, state vectors and gradient vectors are treated as column vectors when they are pre-multiplied by matrices.

[20]Notice that provided $\check{C}$ is compact, we can use (56) to specify a compact set that contains the entire family of minimizers for each $\check{c}$ in $\check{C}$ and a given $\check{x}$.

[21]See Ekeland and Turnbull (1983) for a discussion of continuous time, deterministic control problems when the set of minimizers is not convex. They show that sometimes it is optimal to *chatter* between different controls as a way to imitate convexification in continuous time.

[22]Apply Theorem 2 of Fan (1953) to $-\chi(\cdot, \cdot, \check{x})$. This theorem does not require compactness of the choice set for $\check{h}$, only of the choice set for $\check{c}$. The theorem also does not require attainment when optimization is over the noncompact choice set. In our application, we can verify attainment directly.

The value function V solves:

$$\begin{aligned}\delta V(\check{x}) &= \max_{\check{c}\in\check{C}} \min_{\check{h}} U(\check{c},\check{x}) + \frac{\theta}{2}\check{h}\cdot\check{h} + \left[\mu(\check{c},\check{x}) + \sigma(\check{c},\check{x})\check{h}\right]\cdot V_x(\check{x}) \\ &\quad + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})' V_{xx}(\check{x})\sigma(\check{c},\check{x})\right]\end{aligned}$$

Associated with this value function are the first-order conditions for the controls:

$$\begin{aligned}\theta\check{h} + \sigma(\check{c},\check{x})'\cdot V_x(\check{x}) &= 0 \\ \frac{\partial}{\partial \check{c}}\left(U(\check{c},\check{x}) + \left[\mu(\check{c},\check{x}) + \sigma(\check{c},\check{x})\check{h}\right]\cdot V_x(\check{x}) + \frac{1}{2}\text{trace}\left[\sigma(\check{c},\check{x})' V_{xx}(\check{x})\sigma(\check{c},\check{x})\right]\right) &= 0.\end{aligned}$$

Solving these first-order conditions gives the control laws $h_t = \alpha(x_t)$ and $c_t = \alpha_c(x_t)$. Define μ^* and σ^* such that the states evolve according to

$$dx_t = \mu^*(x_t)dt + \sigma^*(x_t)dB_t$$

after the two optimal controls are imposed. Associated with this recursive representation are processes h and c that can also be depicted as functions of the history of the underlying Brownian motion B.

When the Bellman-Isaacs condition is satisfied, Fleming and Souganidis (1989) provide a formal justification for an equivalent date zero Stackelberg solution in which the minimizing agent announces a decision process $\{h_t : t \geq 0\}$ and the maximizing agent reacts by maximizing with respect to $\{c_t : t \geq 0\}$. We seek a recursive representation of this solution by using a big X, little x formulation. Posit a worst-case process for X_t of the form:

$$dX_t = \mu^*(X_t)dt + \sigma^*(X_t)\left[\alpha_h(X_t)dt + dB_t\right].$$

This big X process is designed so that it produces the same process for $h_t = \alpha_h(X_t)$ that is implied by the Markov perfect equilibrium associated with the value function V when $X_0 = x_0$.

The big X process cannot be influenced by the maximizing agent, but little x can:

$$dx_t = \mu(c_t, x_t)dt + \sigma(c_t, x_t)\left[\alpha_h(X_t)dt + dB_t\right].$$

Combining the two state evolution equations, we have a Markov control problem faced by the maximizing agent. It gives rise to a value function N satisfying a HJB equation:

$$\begin{aligned}\delta N(\check{x},\check{X}) &= \max_{\check{c}\in\check{C}} U(\check{c},\check{x}) + \mu(\check{c},\check{x})\cdot N_x(\check{x},\check{X}) + \mu^*(\check{x})\cdot N_X(\check{X},\check{X}) \\ &\quad + \frac{1}{2}\text{trace}\left(\begin{bmatrix}\sigma(\check{c},\check{x})' & \sigma^*(\check{X})'\end{bmatrix}\begin{bmatrix}N_{xx}(\check{x},\check{X}) & N_{xX}(\check{x},\check{X}) \\ N_{Xx}(\check{x},\check{X}) & N_{XX}(\check{x},\check{X})\end{bmatrix}\begin{bmatrix}\sigma(\check{c},\check{x}) \\ \sigma^*(\check{X})\end{bmatrix}\right) \quad (57) \text{ eq:stack} \\ &\quad + \alpha_h(\check{X})\cdot\sigma(\check{c},\check{x})' N_x(\check{x},\check{X}) + \alpha_h(\check{X})\cdot\sigma^*(\check{X})' N_X(\check{x},\check{X}) \\ &\quad + \frac{\theta}{2}\alpha_h(\check{X})\cdot\alpha_h(\check{X}).\end{aligned}$$

We want the outcome of this optimization problem to produce the same stochastic process for c (c_t as a function of current and past values of the Brownian motion B_t) provided that $X_0 = x_0$. For this to happen, the value functions V and N must be closely related. Specifically,

$$\begin{aligned} N_x(\check{x}, \check{X})|_{\check{X}=\check{x}} &= V_x(\check{x}) \\ N_X(\check{x}, \check{X})|_{\check{X}=\check{x}} &= 0. \end{aligned} \tag{58}$$ eq:first

The first restriction equates the co-state on little x with the implied co-state from the Markov perfect equilibrium along the equilibrium trajectory. The second restriction implies that the co-state vector for big X is zero along this same trajectory.

These restrictions on the first derivative, imply restrictions on the second derivative. Consider a perturbation of the form:

$$\check{x} + \mathbf{r}\nu, \quad \check{X} + \mathbf{r}\nu$$

for some scalar $\mathbf{r}$ and some direction ν. The directions that interest us are those in the range of $\sigma^*(\check{X})$, which are the directions that the Brownian motion can move the state to. Since (58) holds,

$$\begin{aligned} N_{xx}(\check{x}, \check{X})\nu + N_{xX}(\check{x}, \check{X})\nu|_{\check{X}=\check{x}} &= V_{xx}(\check{x})\nu \\ N_{Xx}(\check{x}, \check{X})\nu + N_{XX}(\check{x}, \check{X})\nu|_{\check{X}=\check{x}} &= 0. \end{aligned}$$

¿From HJB (57), we could find a control law that expresses $\check{c}$ as a function of $\check{x}$ and $\check{X}$. We are only concerned, however, with $\check{c}$ evaluated in the restricted domain $\check{x} = \check{X}$. Given the presumed restrictions on the first derivative and the derived restrictions on the second derivative, we can show that $\check{c} = \alpha_c(\check{x})$ satisfies the first-order conditions for $\check{c}$ provided on this restricted domain.

E.2 Changing the objective

The value function for a Bayesian problem does not include a penalty term. In the recursive representation of the date zero Stackelberg problem, the penalty term is expressed completely in terms of big X. We now show how to adjust the value function L by solving a Lyapunov equation.

The function that we wish to compute solves:

$$L(\check{X}) = \frac{\theta}{2} E \int_0^\infty \exp(-\delta t)|\alpha_h(X_t)|^2$$

subject to

$$dX_t = \mu^*(X_t)dt + \sigma^*(X_t)\left[\alpha_h(X_t)dt + dB_t\right].$$

where $X_0 = \check{X}$.

The value function L for this problem solves:

$$\begin{aligned}\delta L(\check{X}) &= \frac{\theta}{2}\alpha_h(\check{X})\cdot\alpha_h(\check{X}) + \mu^*(\check{X})\cdot L_X(\check{X}) \\ &+\frac{1}{2}\text{trace}\left[\sigma^*(\check{X})' L_{XX}(\check{X})\sigma^*(\check{X})\right] + \alpha_h(\check{X})\cdot\sigma^*(\check{X})' L_x(\check{X}).\end{aligned} \tag{59}$$ auglyapunov

E.3 Bayesian value function

To construct a Bayesian value function we form:

$$W(\check{x},\check{X}) = N(\check{x},\check{X}) - L(\check{X}).$$

Given equations (57) and (59), the separable structure of W implies that it satisfies the HJB equation:

$$\begin{aligned}\delta W(\check{x},\check{X}) &= \max_{\check{c}\in\check{C}} U(\check{c},\check{x}) + \mu(\check{c},\check{x})\cdot W_x(\check{x},\check{X}) + \mu^*(\check{x})\cdot W_X(\check{X},\check{X}) \\ &+\frac{1}{2}\text{trace}\left(\begin{bmatrix}\sigma(\check{c},\check{x})' & \sigma^*(\check{X})'\end{bmatrix}\begin{bmatrix}W_{xx}(\check{x},\check{X}) & W_{xX}(\check{x},\check{X})\\ W_{Xx}(\check{x},\check{X}) & W_{XX}(\check{x},\check{X})\end{bmatrix}\begin{bmatrix}\sigma(\check{c},\check{x})\\ \sigma^*(\check{X})\end{bmatrix}\right) \\ &+\alpha_h(\check{X})\cdot\sigma(\check{c},\check{x})' W_x(\check{x},\check{X}) + \alpha_h(\check{X})\cdot\sigma^*(\check{X})' W_X(\check{x},\check{X})\end{aligned}$$

Then $\check{z}W(\check{x},\check{X})$ the value function for the stochastic control problem:

$$\check{z}W(\check{x},\check{X}) = E\int_0^\infty \exp(-\delta t) z_t U(c_t,x_t)dt$$

and evolution:

$$\begin{aligned}dx_t &= \mu(c_t,x_t)dt + \sigma(c_t,x_t)dB_t \\ dz_t &= z_t\alpha_h(X_t)dB_t \\ dX_t &= \mu^*(X_t)dt + \sigma^*(X_t)dB_t\end{aligned}$$

where $z_0 = \check{z}$, $x_0 = \check{x}$ and $X_0 = \check{X}$. To interpret the nonnegative z as inducing a change in probability, we initialize z_0 at unity.

Also, $W(\check{x},\check{X},\theta)$ is the value function for a control problem with discounted objective:

$$W(\check{x},\check{X}) = \max_{c\in C} E\int_0^\infty \exp(-\delta t) U(c_t,x_t)dt$$

and evolution:

$$\begin{aligned}dx_t &= \mu(c_t,x_t)dt + \sigma(c_t,x_t)\left[\alpha_h(X_t)dt + d\tilde{B}_t\right] \\ dX_t &= \mu^*(X_t)dt + \sigma^*(X_t)\left[\alpha_h(X_t)dt + d\tilde{B}_t\right].\end{aligned}$$

This value function is constructed using a perturbed specification where a Brownian increment dB_t is replaced by an increment $\alpha_h(X_t)dt + d\tilde{B}_t$ with a drift distortion that depends only on the uncontrollable state X. This perturbation is justified via the Girsanov Theorem, provided that we entertain a weak solution to the stochastic differential equation governing the state evolution equation.

References

ahs2003 Anderson, E., L. P. Hansen, and T. Sargent (2003, March). A quartet of semi-groups for model specification, robustness, prices of risk, and model detection. *Journal of the European Economic Association 1*, 68–123.

raujosandroni Araujo, A. and A. Sandroni (1999). On the convergence to homogeneous expectations when markets are complete. *Econometrica 67*(3), 663–672.

basarbernhard Basar, T. and P. Bernhard (1995). *H_∞-Optimal Control and Related Minimax Design Problems*. Boston: Birkhauser.

ackwelldubins Blackwell, D. and L. Dubins (1962). Merging of opinions with increasing information. *Annals of Mathematical Statistics 38*, 882–886.

kwellgirshick Blackwell, D. and M. Girshick (1954). *Theory of Games and Statistical Decisions*. New York: Wiley.

chamberlain Chamberlain, G. (2000). Econometric applications of maxmin expected utility theory. *Journal of Applied Econometrics 15*, 625–644.

duffieepstein Duffie, D. and L. G. Epstein (1992). Stochastic differential utility. *Econometrica 60*(2), 353–394.

elandturnbull Ekeland, I. and T. Turnbull (1983). *Infinite-Dimensional Optimization and Convexity*. Chicago Lectures in Mathematics. Chicago: The University of Chicago Press.

teinschneider Epstein, L. and M. Schneider (2001). Recursive multiple-priors. unpublished.

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

fan1 Fan, K. (1952). Fixed-point and minimax theorems in locally convex topological linear sapces. *Proceedings of the National Academy of Sciences 38*, 121–126.

fan2 Fan, K. (1953). Minimax theorems. *Proceedings of the National Academy of Sciences 39*, 42–47.

flemingsoner Fleming, W. H. and H. M. Soner (1993). *Controlled Markov Processes and Viscosity Solutions*. Applications of Mathematics. New York: Springer-Verlag.

ingsouganidis Fleming, W. H. and P. E. Souganidis (1989). On the existence of value functions of two-player, zero sum stochastic differential games. *Indiana University Mathematics Journal 38*, 293 – 314.

follmer Föllmer, H. (1985). An entropy approach to the time reversal of diffusion processes. In M. Metivier and E. Pardoux (Eds.), *Stochastic Differential Systems*, Volume 69 of *Lecture Notes in Control and Information Sciences*, pp. 156–163. Springer-Verlag.

boaschmeidler Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*, 141–153.

argentieeetac Hansen, L. P. and T. Sargent (1995, May). Discounted linear exponential quadratic gaussian control. *IEEE Transactions on Automatic Control 40*(5), 968–971.

robustpi Hansen, L. P., T. Sargent, and T. Tallarini (1999). Robust permanent income and pricing. *Review of Economic Studies 66*, 873–907.

ensargentking Hansen, L. P. and T. J. Sargent (2003, April). Robust control of forward-looking models. *Journal of Monetary Economics 50*(3), 581 – 604.

ensargentbook Hansen, L. P. and T. J. Sargent (2005). *Misspecification*. Princeton.

jackson Jackson, M. O., E. Kalai, and R. Smordoninsky (1999, July). Bayesian representation of stochastic processes under learning: De Finetti revisited. *Econometrica 67*(4), 875–893.

jacobson Jacobson, D. H. (1973). Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Transactions on Automatic Control 18*, 124–131.

jamesperturb James, M. R. (1992). Asymptotic analysis of nonlinear stochastic risk sensitive control and differential games. *Mathematics of Control, Signals, and Systems 5*, 401–417.

kls Kabanov, J. M., R. S. Lipcer, and A. N. Sirjaev (1979). Absolute continuity and singularity of locally absolutely continuous probability distributions. I. *Math. USSR Sbornik 35*, 631 – 696.

kalai Kalai, E. and E. Lerner (1993). Rational learning leads to Nash equilibrium. *Econometrica 61*(5), 1019–1045.

aratzasshreve Karatzas, I. and S. E. Shreve (1991). *Brownian Motion and Stochastic Calculus* (Second ed.). New York: Springer-Verlag.

kunita Kunita, H. (1969). Absolute continuity of Markov processes and generators. *Nagoya Mathematical Journal 36*, 1–26.

ptsershiryaev Liptser, R. S. and A. N. Shiryaev (2000). *Statistics of Random Processes* (Second ed.), Volume I. General Theory of *Applications of Mathematics*. Berlin: Springer.

luenberger Luenberger, D. G. (1969). *Optimization by Vector Space Methods*. New York: Wiley.

james2 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.

revuz Revuz, D. and M. Yor (1994). *Continuous Martingales and Brownian Motion* (Second ed.). New York: Springer-Verlag.

sandroni2000 Sandroni, A. (2000). Do markets favor agents able to make accurate predictions? *Econometrica 68*(6), 1303–1341.

whittle Whittle, P. (1981). Risk-sensitive linear/quadratic/gaussian control. *Advances in Applied Probability 13*, 764–777.