Robust Control and Model Uncertainty

Lars Peter Hansen   Thomas J. Sargent

January 22, 2001

I Introduction

This paper describes links between the max-min expected utility theory of Gilboa and Schmeidler (1989) and the applications of robust control theory proposed by Anderson, Hansen and Sargent (2000) and Dupuis, James and Petersen (1998).[1] The max-min expected utility theory represents uncertainty aversion with preference orderings over decisions c and states x, for example, of the form

$$\inf_{Q \in \mathcal{Q}} E_Q \left[\int_0^\infty \exp(-\delta t) U(c_t, x_t) \, dt \right] \tag{1}$$

where $\mathcal{Q}$ is a set of measures over c, x, and δ a discount rate. Gilboa and Schmeidler's theory leaves open how to specify the set $\mathcal{Q}$ in particular applications.[2]

Criteria like (1) also appear as objective functions in robust control theory. Robust control theory specifies $\mathcal{Q}$ by taking a single 'approximating model' and statistically perturbing it;

$\mathcal{Q}$ is typically parameterized only implicitly, through a positive penalty variable θ. This paper describes how to transform that 'penalty problem' into a closely related 'constraint problem' like (1). These two formulations differ in subtle ways but are connected via the Lagrange multiplier theorem. The implicit preference orderings differ but imply the same decisions. Both preferences are recursive, and therefore both are time consistent. However, time consistency for the constraint specification requires that we introduce a new endogenous state variable to restrict how probability distortions are reconsidered at future dates. To facilitate comparisons to Anderson et al. (2000) and Chen and Epstein (2000), we cast our discussion within continuous-time diffusion models.

II A Benchmark Resource Allocation Problem

We first pose a discounted, infinite time optimal resource allocation problem without regard to robustness. Let $\{B_t : t \geq 0\}$ denote a d-dimensional, standard Brownian motion on an underlying probability space $(\Omega, \mathcal{F}, P)$. Let $\{\mathcal{F}_t : t \geq 0\}$ denote the completion of the filtration generated by this Brownian motion. The actions of the decision-maker form a stochastic process $\{c_t : t \geq 0\}$ that is progressively measurable. Let U denote an instantaneous utility function, and write the discounted objective as $\sup_{c \in C} E\left[\int_0^\infty \exp(-\delta t) U(c_t, x_t) d\,t\right]$ subject to:

$$dx_t = \mu(c_t, x_t)dt + \sigma(c_t, x_t)dB_t \tag{2}$$

where x_0 is a given initial condition and C is a set of admissible control processes. We use P to denote the stochastic process for x_t generated by (2). Equation (2) will be the 'approximating model' of later sections, to which all other models in $\mathcal{Q}$ are perturbations.

We restrict μ and σ so that any progressively measurable control c in C implies a progressively measurable state vector process x. We assume throughout that the objective for the control problem without reference to robustness has a finite upper bound.

III Model Misspecification

The decision maker treats (2) as an approximation by taking into account a class of alternative models that are statistically difficult to distinguish from (2). To construct a perturbed model we replace B_t in (2) by $\hat{B}_t + \int_0^t h_s ds$ where h is progressively measurable and $\{\hat{B}_t\}$ is a Brownian motion. Then we can write the distorted stochastic evolution in continuous time as $dx_t = \mu(c_t, x_t)dt + \sigma(c_t, x_t)(h_t dt + d\hat{B}_t)$ under the Brownian motion probability specification.

A Changes in Measure

The process h is used as device to transform the probability distribution P on $(\Omega, \mathcal{F})$ into a new distribution Q that is absolutely continuous with respect to P. An absolutely continuous change in measure for a stochastic process can be represented in terms of a nonnegative martingale. Let Q denote a probability distribution that is absolutely continuous with respect to P. Associated with Q is a family of expectation operators applied to random variables that are $\mathcal{F}_t$ measurable for each t. Thus we can write $E_Q g_t = E_P g_t q_t$ for any bounded g_t

that is $\mathcal{F}_t$ measurable and some nonnegative random variable q_t that is $\mathcal{F}_t$ measurable. The random variable q_t is called a Radon-Nikodym derivative. In our setting, we use the Girsanov Theorem to depict q_t as $q_t = \exp\left[\int_0^t h_\tau \cdot d\hat{B}_\tau - \int_0^t \frac{|h_\tau|^2}{2} dt\right]$. We use this representation to justify our use of h to parameterize absolutely continuous changes of measure.[3] When h is zero we revert to the benchmark control problem.

B Relative Entropy of a Stochastic Process

Consider a scalar stochastic process $\{g_t\}$ that is progressively measurable. This process is a random variable on a product space. Form $\Omega^* = \Omega \times \mathbb{R}^+$ where $\mathbb{R}^+$ is the nonnegative real line; form the corresponding sigma algebra $\mathcal{F}^*$ 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 P^* as the product measure $P \times M$ where M is exponentially distributed with density $\delta \exp(-\delta t)$. We let E^* denote the expectation operator on the product space. The E^* expectation of the stochastic process $\{g_t\}$ is by construction $E^*(g) = \delta \int_0^\infty \exp(-\delta t) E(g_t) dt$.

We extend this construction by using the probability measure Q. Form $Q^* = Q \times M$. The process $\{q_t\}$ is a Radon-Nikodym derivative for Q^* with respect to P^*: $E_Q^*(g) = \delta \int_0^\infty \exp(-\delta t) E(q_t g_t) dt$. The Q^* can be used to evaluate discounted expected utility under an absolutely continuous change in measure.

We measure the discrepancy between the distributions of P and Q as the *relative entropy* between Q^* and P^*: $\mathcal{R}(Q) = \delta \int_0^\infty \exp(-\delta t) E_Q \left(\log q_t\right) dt = \int_0^\infty \exp(-\delta \tau) E_Q \left(\frac{|h_\tau|^2}{2}\right) d\tau$. Relative entropy is convex in the measure Q^* (*e.g.* see Dupuis and Ellis (1997)). Relative entropy is nonnegative and zero only when the probability distributions P^* and Q^* agree. This is true only when the process h is zero.

IV Two Robust Control Problems

We study the relationship between two robust control problems. Let E_Q denote the mathematical expectation taken with respect to the stochastic process $\{B_t : t \geq 0\}$ where $dB_t = d\hat{B}_t + h_t dt$ and $\{\hat{B}_t : t \geq 0\}$ is a Brownian motion under both P and Q. Thus we parameterize Q by the choice of drift distortion $\{h_t\}$, and use the state evolution equation:

$$dx_t = \mu(c_t, x_t)dt + \sigma(c_t, x_t)dB_t. \tag{3}$$

We define two control problems. A **multiplier** robust control problem is $\sup_{c\in C} \inf_Q E_Q\left[\int_0^\infty \exp(-\delta t)U(c_t, x_t)d\,t\right] + \theta\mathcal{R}(Q)$ subject to (3). A **constraint** robust control problem is $\sup_{c\in C} \inf_Q E_Q\left[\int_0^\infty \exp(-\delta t)U(c_t, x_t)d\,t\right]$subject to (3) and $\mathcal{R}(Q) \leq \eta$. Note that $\mathcal{R}(Q) \leq \eta$ is a single intertemporal constraint on the entire path of distortions h.

These two problems are closely related. We can interpret the robustness parameter θ in the first problem as an implied Lagrange multiplier on the specification-error constraint $\mathcal{R}(Q) \leq \eta$.[4] Use θ to index a family of multiplier robust control problems and η to index a family of constraint robust control problems. Because not all values of θ are admissible, we consider only nonnegative values of θ for which it is feasible to make the objective function greater than $-\infty$. Call the closure of this set Θ. In Hansen, Sargent, Turmuhambetova and

Williams (2001) we provide assumptions and a proof for:

Claim IV.1. *Suppose that for $\eta = \eta^*$, c^* and Q^* solve the constraint robust control problem. There exists a $\theta^* \in \Theta$ such that the multiplier and constraint robust control problem have the same solution.*

To construct the multiplier, let $J(c, \eta)$ satisfy $J(c, \eta) = \inf_Q E_Q \left[\int_0^\infty \exp(-\delta t) U(c_t, x_t) dt \right]$, subject to $\mathcal{R}(Q) \leq \eta$ and $J^*(\eta) = \sup_{c \in C} J(c, \eta)$. As argued by Luenberger (1969), $J(c, \eta)$ is decreasing and convex in η. These same properties carry over to the optimized (over c) function J^*. Given η^*, we let θ^* be the negative of the slope of the subgradient of J^* at η^*, i.e., θ^* is the absolute value of the slope of a line tangent to J^* at η^*.

Hansen et al. (2001) also establish:

Claim IV.2. *Suppose J^* is strictly decreasing, θ^* is in the interior of Θ, and that there exists a solution c^* and Q^* to the multiplier robust control problem. Then that c^* also solves the constraint robust control problem for $\eta = \eta^* = \mathcal{R}(Q^*)$.*

Claims IV.1 and IV.2 are observational equivalence results because they describe how the multiplier and constraint robust control problems give rise to the same decisions. By adapting arguments in Hansen and Sargent (1995) and Anderson et al. (2000), it can be shown that the multiplier robust control problem has the same solution as a recursive risk-sensitive control problem, where $-\theta^{-1}$ is the risk-sensitivity parameter.[5] Claims IV.1 and IV.2 thus link a risk-sensitive control problem to the constraint robust control problem.

V Recursivity of the Multiplier Formulation

The multiplier robust control problem can be represented as $\sup_c \inf_h \hat{E} \int_0^\infty \exp(-\delta t) \left[U(c_t, x_t) + \frac{\theta}{2}(h_t \cdot h_t)\right] dt$ subject to $dx_t = \mu(c_t, x_t)dt + \sigma(c_t, x_t)(h_t dt + d\hat{B}_t)$. We can view h as a second control process in a two-player zero-sum game. Given h we can fix the distribution for $\hat{B}$ as a multivariate standard Brownian motion. Then there is a single probability distribution in play and we use the notation $\hat{E}$ to denote the associated expectation operator. Fleming and Souganidis (1989) tell how a Bellman-Isaacs condition justifies a recursive solution by relating a solution to a date zero *commitment* game to a Markov perfect game in which the decision rules of both agents are functions of the state vector x_t. The Bellman-Isaacs condition is:

Assumption V.1. *There exists a value function V such that:*

$$
\begin{aligned}
\delta V &= \max_c \min_h U(c, x) + \frac{\theta}{2} h \cdot h + [\mu(c, x) + \sigma(c, x)h] \cdot \frac{\partial V(x)}{\partial x} \\
&\quad + \text{trace}\left[\sigma(c, x)' \frac{\partial^2 V(x)}{\partial x \partial x'} \sigma(c, x)\right] \\
&= \min_h \max_c U(c, x) + \frac{\theta}{2} h \cdot h + [\mu(c, x) + \sigma(c, x)h] \cdot \frac{\partial V(x)}{\partial x} \\
&\quad + \text{trace}\left[\sigma(c, x)' \frac{\partial^2 V(x)}{\partial x \partial x'} \sigma(c, x)\right].
\end{aligned}
$$

The Bellman-Isaacs condition defines a Bellman equation for a two-player zero-sum game in which both players decide at time 0 or recursively. The associated decision rules for c and h also solve our two robust control problems.

VI Two Preference Orderings

While the Lagrange multiplier theorem links the two robust control problems, the implied preference orders differ. But they are related at the common solution to both problems, where their indifference curves are tangent.

A Preference Orderings

To construct two preference orderings, we assume an endogenous state vector s_t:

$$ds_t = \mu_s(s_t, c_t)dt. \tag{4}$$

where this differential equation can be solved uniquely for s_t given s_0 and process $\{c_s : 0 \leq s < t\}$. We assume that the solution is a progressively measurable process $\{s_t : t \geq 0\}$. We think of s_t as an endogenous component of the state vector x_t. We can use s_t to make preferences nonseparable over time as in models with habit persistence. We use the felicity function $u(s_t, c_t)$ to represent preferences that are additively separable in (s_t, c_t).

We define preference orders for times $\tau \geq 0$ in terms of two functions, $D_\tau(c, s_\tau)$, $R_\tau(Q)$. First, define $D_\tau(c, s_\tau) = \int_\tau^\infty \exp(-\delta t) u(s_{t+\tau}, c_{t+\tau}) dt$ where s_τ is the date τ initial condition for differential equation (4). The impact of consumption between dates 0 and τ is captured by the state variable s_τ.

Next, define a time τ model discrepancy measure $\mathcal{R}_\tau(Q) = \delta \int_0^\infty \exp(-\delta t) E_Q \left(\log q_{t+\tau} - \log q_\tau | \mathcal{F}_\tau\right) dt$. The local evolution of $\mathcal{R}(Q)$ is $d\mathcal{R}_t(Q) =$

$\left[-\frac{|h_t|^2}{2} + \delta \mathcal{R}_t(Q)\right] dt$ with initial condition: $\mathcal{R}_0(Q) = \mathcal{R}(Q)$. We use $D_\tau(c, s_\tau)$ to represent both preference specifications at τ, and use $R_\tau(Q)$ to help us represent preferences under the constraint specification.

For fixed θ, we represent the date τ *multiplier* preferences using the valuation function $\hat{W}_\tau(c;\theta) = \inf_Q E_Q\left[D_\tau(c, s_\tau)|\mathcal{F}_\tau\right] + \theta \mathcal{R}_\tau(Q)$. For a nonnegative r_τ that is $\mathcal{F}_\tau$ measurable, we represent the time τ *constraint* preferences in terms of the valuation function $W_\tau(c; r_\tau) = \inf_{\mathcal{R}_\tau(Q)\leq r_\tau} E_Q\left[D_\tau(c, s_\tau)|\mathcal{F}_\tau\right]$. For convenience, denote the time 0 version $W_0(c, r_0) = W(c, \eta)$ and the time 0 version $\hat{W}_0(c,\theta) = \hat{W}(c,\theta)$.

We define preference orderings as follows. For any two progressively measurable c and c^*, $c^* \succeq_\eta c$ if $W(c^*;\eta) \geq W(c;\eta)$. For any two progressively measurable c and c^*, $c^* \hat{\succeq}_\theta c$ if $\hat{W}(c^*;\theta) \geq \hat{W}(c;\theta)$. We would use analogous definitions for time τ versions of the preference orderings.

The multiplier preference ordering coincides with a recursive, risk sensitive preference ordering provided that $\theta > 0$.[6]

B Relation between the Preference Orders

The two time 0 preference orderings differ. Furthermore, given η, there exists no θ that makes the two preference orderings agree. However, the Lagrange Multiplier Theorem delivers a weaker result that is very useful to us. While globally the preference orderings differ, indifference curves that pass through the solution c^* to the optimal resource allocation problem are tangent.

Use the Lagrange Multiplier Theorem to write $W(c^*;\eta^*) = \max_\theta \inf_Q E_Q D(c^*) + \theta\left[\mathcal{R}(Q) - \eta^*\right]$, and let θ^* denote the maximizing value of θ, which we assume to be strictly positive. Suppose that $c^* \succeq_{\eta^*} c$. Then $\hat{W}(c;\theta^*) - \theta^*\eta^* \leq W(c;\eta^*) \leq W(c^*;\eta^*) = \hat{W}(c^*;\theta^*) - \theta^*\eta^*$. Thus $c^* \hat{\succeq}_{\theta^*} c$.

The observational equivalence results from claims IV.1 and IV.2 apply to consumption profile c^*. At this point, the indifference curves are tangent, implying that they are supported by the same prices. Observational equivalence claims made by econometricians typically refer to equilibrium trajectories and not to off-equilibrium aspects of the preference orders.

VII Recursivity of the Preference Orderings

To study time consistency, we describe the relation between the time zero and time $\tau > 0$ valuation functions that define preference orders. At date τ, some information has been realized and some consumption has taken place. Our preference orderings focus the attention of the decision-maker on subsequent consumption in states that can be realized given current information. These considerations underlie our use of D_τ and $\mathcal{R}_\tau$ to depict $W_\tau(c,\theta)$ and $\hat{W}_\tau(c, r_\tau)$. The function D_τ reflects a change in vantage point as time passes. Except through s_τ, the function D_τ depends only on the consumption process from date τ forward.

In addition, at date τ the decision maker focuses on states that can be realized from date τ forward. Expectations used to average over states are conditioned on date τ information. In this context, while conditioning on time τ information, it would be inappropriate to constrain probabilities using only date zero relative entropy. Imposing a date zero relative entropy constraint at date τ would introduce a temporal inconsistency by letting the minimizing agent put no probability distortions at dates that have already occurred and in states that at date τ cannot be realized. Instead, we make the date τ decision-maker explore only probability distortions that alter his preferences from date τ forward. This leads us to use $\mathcal{R}_\tau$ as a conditional counterpart to our relative entropy measure.

Our entropy measure has a recursive structure. Date zero relative entropy is easily constructed from the conditional relative entropies in future time periods. We can write:

$$\mathcal{R}(Q) = E_Q \left[\int_0^\tau \exp(-\delta t) \frac{|h_t|^2}{2} dt + \exp(-\delta \tau) \mathcal{R}_\tau(Q) \right] \tag{5}$$

The recursive structure of the multiplier preferences follows from this representation. In effect the date zero valuation function $\hat{W}$ can be separated by disjoint date τ events and depicted as $\hat{W}(c;\theta) = \inf_{\{h_t: 0 \le t < \tau\}} \hat{E} \left(\int_0^\tau \exp(-\delta t) \left[U(c_t, s_t) + \theta \frac{|h_t|^2}{2} \right] dt + \hat{W}_\tau(c;\theta) \right)$ subject to

$$\begin{aligned} dB_t &= d\hat{B}_t + h_t dt && (6) \\ ds_t &= \mu_s(s_t, c_t) dt. && (7) \end{aligned}$$

The constraint preferences at time τ make the decision-maker explore changes in probability distributions from date τ forward. We also want to exclude the possibility of changing the probabilities of events known in previous dates and of events known not to occur. For the date zero constraint preferences, given c we can find an $\tilde{h}$ process used to construct $W(c, \eta)$.

Associated with this $\tilde{h}$ process, we can compute the time τ conditional relative entropy $\mathcal{R}_\tau(\tilde{Q})$. Thus, implicit in the construction of the valuation function $W(c,\eta)$ is a partition of relative entropy over time and across states as in (5). At date τ we ask the decision-maker to explore only changes in beliefs that effect outcomes that can be realized in the future. That is, we impose the constraint $\mathcal{R}_\tau(Q) \leq r_\tau$ for $r_\tau = \mathcal{R}_\tau(\tilde{Q})$ along with fixing $\tilde{h}_t$ for $0 \leq t < \tau$. Notice that with this constraint imposed, $\mathcal{R}(Q) \leq \mathcal{R}(\tilde{Q})$ so that we continue to satisfy our date zero relative entropy constraint. We tie the hands of the date τ decision-maker to inherit how conditional relative entropy is to be allocated across states that are realized at date τ. (Chen and Epstein (2000) avoid this extra hand-tying by imposing separate constraints on h for every date and state.) We can write the valuation function for the constrained problem recursively as $W(c,\eta) = \inf_{\{h_t:0\leq t<\tau\}} \hat{E}\int_0^\tau \exp(-\delta t)U(c_t,s_t)dt + \hat{E}W_\tau(c,r_\tau)$ subject to (6), (7) and $r_\tau \geq 0$, where r_τ solves $dr_t = \left(\delta r_t - \frac{|h_t|^2}{2}\right)dt$ for $0 \leq t < \tau$ with initial condition $r_0 = \eta$.

VIII Concluding Remarks

Empirical work in macroeconomics and finance typically assumes a unique and explicitly specified dynamic statistical model. To use Gilboa and Schmeidler (1989)'s multiple-model expected utility theory, we have turned to robust control theory for a parsimonious (one parameter) set of alternative models with rich alternative dynamics. Those alternative models come from perturbing the decision maker's approximating model to allow its shocks to feed back on state variables arbitrarily. This allows the approximating model to miss functional forms, the serial correlation of shocks and exogenous variables, and how those exogenous variables impinge on endogenous state variables. Anderson et al. (2000) show how the multiplier parameter in the robust control problems indexes a set of perturbed models that is difficult to distinguish statistically from the approximating model given a sample of T time-series observations.

References

Anderson, Evan, Lars Peter Hansen, and Thomas Sargent, "Robustness, Detection and the Price of Risk," March 2000. Mimeo.

Chen, Z. and L. G. Epstein, "Ambiguity, Risk and Asset Returns in Continuous Time," 2000. unpublished manuscript.

Duffie, Darrell and Larry G. Epstein, "Stochastic Differential Utility," *Econometrica*, 1992, *60* (2), 353–394.

Dupuis, P. and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations* Wiley Series in Probability and Statistics, New York: John Wiley and Sons, 1997.

Dupuis, Paul, Matthew R. James, and Ian Petersen, "Robust Properties of Risk Sensitive Control," LCDS 98-15, Brown University 1998.

Epstein, L. and S. Zin, "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 1989, *57*, 937–969.

Fleming, W. H. and P. E. Souganidis, "On the Existence of Value Functions of Two-Player, Zero Sum Stochastic Differential Games," *Indiana University Mathematics Journal*, 1989, *38*, 293 – 314.

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

Hansen, L. P. and T. J. Sargent, *Elements of robust control and filtering for* *macroeconomics*, To be decided, 2001.

– , – , G. A. Turmuhambetova, and N. Williams, "Robust Control and Model Uncertainty," 2001. Mimeo, University of Chicago.

Hansen, Lars Peter and Thomas Sargent, "Discounted Linear Exponential Quadratic Gaussian Control," *IEEE Transactions on Automatic Control*, May 1995, *40* (5), 968–971.

– , – , and Thomas Tallarini, "Robust Permanent Income and Pricing," *Review of Economic Studies*, 1999, *66*, 873–907.

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

Peterson, I. R., M. R. James, and P. Dupuis, "Minimax Optimal Control of Stochastic Uncertain Systems with Relative Entropy Constraints," *IEEE Transactions on Automatic Control*, 2000, *45*, 398–412.

Notes

[1]Lars Peter Hansen, Department of Economics, University of Chicago, 1126 E. 59th St., Chicago, Illinois 60637; Thomas J. Sargent, Hoover Institution and Department of Economics, Stanford University, Stanford, California 94305.

[2]This paper summarizes detailed arguments in Hansen et al. (2001).

[3]Perturbations that are not absolutely continuous are easy to detect statistically, which is the reason that Anderson et al. (2000) impose absolute continuity on the perturbations.

[4] This connection has been explored in informally in Hansen, Sargent and Tallarini (1999) and formally in Hansen and Sargent (2001) in the context of linear-quadratic control problem. We mimic arguments in Peterson, James and Dupuis (2000) and Luenberger (1969).

[5]Risk-sensitive control theory makes decision rules more responsive to risk by making an exponential adjustment to the objective of the decision-maker in the same way used by Epstein and Zin (1989) and Duffie and Epstein (1992). Hansen and Sargent (1995) and Anderson et al. (2000) show how risk-sensitive control theory can be motivated through recursive utility theory.

[6]Under the Brownian motion information structure, these multiplier preferences coincide with a special case of stochastic differential utility studied by Duffie and Epstein (1992)